TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

WILLIAM C. GRAUSTEIN

EINAR HILLE

C. C. MACDUFFEE

VOLUME 41

JANUARY TO JUNE, 1937

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1937

Chester C. Corbin Library Fund
Bound August, 1937

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

TABLE OF CONTENTS

VOLUME 41, JANUARY TO JUNE, 1937

ON PHRAGMÉN-LINDELÖF'S PRINCIPLE*

BY

LARS V. AHLFORS

In the first part of this paper we give a proof of Phragmén-Lindelöf's now classical principle,† which is simpler and yields more detailed information than any of the proofs hitherto known. Our procedure consists in proving a theorem in finite terms, similar to Hadamard's three circles theorem, from which the asymptotic statement is shown to follow by a very simple and transparent reasoning. A certain symmetry in the result is obtained by allowing the functions considered to have two possible singularities, one at 0 and one at ∞. The ultimate theorem is the sharpest possible and contains all previous results, including those of the brothers Nevanlinna.‡

In Part II we generalize Phragmén-Lindelöf's principle to harmonic functions of n variables. The methods of Part I are seen to carry over without any difficulties. The result is particularly interesting in so far as the symmetry of the two-dimensional case is not maintained, the extremal functions corresponding to the two singularities being now essentially different.

PART I

1. Let $f(s)=f(x+iy)$ be analytic in the strip T: $-\pi/2 \leqq y \leqq \pi/2$, and suppose further that $|f(x \pm i\pi/2| \leqq 1$ for every x. Under these fundamental assumptions two alternatives are possible:

(a) The inequality $|f(s)| \leqq 1$ holds for all s in T. (The situation is then governed by theorems such as the lemmas of Schwarz and Julia, etc., and shall not concern us further.)

(b) The set of points s in T with $|f(s)| > 1$ is not void.

In the sequel we shall always suppose that the second alternative takes place.

Some of our results are stated for a finite interval (x_1, x_2). In that case it will be noted that the arguments and results are not based on any assump-

* Presented to the Society, April 10, 1936; received by the editors March 24, 1936.

† E. Phragmén and E. Lindelöf, *Sur une extension d'un principe classique de l'analyse et sur quelque propriétés des fonctions monogènes dans le voisinage d'un point singulier*, Acta Mathematica, vol. 31 (1908), pp. 381–406.

‡ F. and R. Nevanlinna, *Über die Eigenschaften analytischer Funktionen in der Umgebung einer singularen Stelle oder Linie*, Acta Societatis Scientiarum Fennicae, vol. 50 (1922), No. 5.

R. Nevanlinna, *Über die Eigenschaften meromorpher Funktionen in einem Winkelraum*, Acta Societatis Scientiarum Fennicae, vol. 50 (1922), No. 12.

tions regarding the function outside of the interval with which we are concerned. For the sake of convenience we shall not further repeat this remark.

Consider now the set of points for which $|f(s)| > 1$. Obviously this set is composed by certain open regions Ω, none of which is finite. We choose arbitrarily a region Ω and propose to study the behavior of $f(s)$ within this region. As a preliminary result we notice that $|f(s)| = 1$ on the boundary of Ω, and that this boundary is made up of analytic arcs, some of which may be segments of the lines $y = \pm\pi/2$. We also remark that the boundary may form an angle, namely, at points where $f'(s) = 0$.

Denote by Δ_t the set in which the line $x = t$ intersects the region Ω. We introduce the notation

$$m(x) = \int_{\Delta_x} \log | f(x + iy) | \cos y \, dy \tag{1}$$

and complete the definition by setting $m(x) = 0$ if Δ_x is void. The function $m(x)$ is evidently non-negative and continuous.*

Differentiating (1) at a point x, where the boundary of Ω has no vertical tangent and forms no angle, we first obtain

$$m'(x) = \int_{\Delta_x} \frac{\partial}{\partial x} \log | f(x + iy) | \cos y \, dy,$$

since $\log |f| = 0$ at the endpoints of the segments forming Δ_x. One more differentiation gives

$$m''(x) = \int_{\Delta_x} \frac{\partial^2}{\partial x^2} \log | f | \cos y \, dy + \frac{\partial}{\partial x} \log | f | \cos y \frac{dy}{dx} \Big|_1^2, \tag{2}$$

where the sign $\big|_1^2$ means that we have to form the sum of the values of the expression under the sign for all upper endpoints of the intervals Δ_x and subtract the corresponding sum for the lower endpoints.

Observing that $\Delta \log |f| = 0$ the integral may be transformed and integrated by parts as follows:

$$\begin{aligned}\int_{\Delta_x} \frac{\partial^2}{\partial x^2} \log | f | \cos y \, dy &= - \int_{\Delta_x} \frac{\partial^2}{\partial y^2} \log | f | \cos y \, dy \\ &= - \int_{\Delta_x} \frac{\partial}{\partial y} \log | f | \sin y \, dy - \frac{\partial}{\partial y} \log | f | \cos y \Big|_1^2 \\ &= \int_{\Delta_x} \log | f | \cos y \, dy - \frac{\partial}{\partial y} \log | f | \cos y \Big|_1^2 .\end{aligned}$$

* The idea of introducing a quantity like $m(x)$ is due to the brothers Nevanlinna (loc. cit.) who consider $\int_{-\pi/2}^{\pi/2} \overset{+}{\log} |f| \cos y \, dy$. This integral evidently represents the sum of the $m(x)$ for all regions Ω and is consequently not less than every single $m(x)$.

Substituting in (2) we finally obtain the formula

$$m''(x) = m(x) + \left(\frac{\partial}{\partial x}\log|f|\frac{dy}{dx} - \frac{\partial}{\partial y}\log|f|\right)\cos y\,\Big|_1^2.$$

Here the terms of the finite sum are all seen to be non-negative, and we obtain

$$m''(x) \geqq m(x). \tag{3}$$

We still have to consider the obviously isolated irregular values where one of the exceptional circumstances occurs. First of all it may happen that a line $x=t$ contains a whole segment lying on the boundary of Ω. It is then easy to prove that the one-sided derivatives $m'(t-0)$ and $m'(t+0)$ exist, and that $m'(t-0)<m'(t+0)$. At a point with vertical tangent and at a corner $m'(x)$ is seen to be continuous, while $m''(x)$ may fail to exist.

These remarks complete the proof of

THEOREM 1. *The function $m(x)$ satisfies the differential inequality*

$$m''(x) \geqq m(x) \tag{3}$$

except at certain isolated points. At these exceptional points $m'(x)$ is continuous or presents a positive jump.

The sign of equality holds if and only if Δ_x coincides with the segment $(-\pi/2, \pi/2)$ or is vacuous.

2. The solutions of the differential equation $\phi''(x)=\phi(x)$ are $\phi(x)=\alpha e^x +\beta e^{-x}$. We are going to show that the curve C: $Y=m(x)$ is convex with respect to this family of functions.

Suppose that the curves $Y=m(x)$ and $Y=\phi(x)$ intersect in two points with the coordinates x_1 and $x_2(>x_1)$, and let these be the nearest points of intersection so that there are no common points between x_1 and x_2. The difference $\omega(x)=m(x)-\phi(x)$ has a constant sign in (x_1, x_2) and satisfies the differential inequality $\omega''(x)\geqq\omega(x)$. This leads to a contradiction if $\omega(x)$ is constantly positive. In fact, if $t_1<t_2<\cdots<t_k$ are the irregular points between x_1 and x_2, we have

$$\begin{aligned}\omega'(x_1+0) + [\omega'(t_1+0) - \omega'(t_1-0)] + \cdots \\ + [\omega'(t_k+0) - \omega'(t_k-0)] - \omega'(x_2-0) \geqq 0,\end{aligned}$$

since all the terms are non-negative. Using the mean-value theorem we can rewrite the expression in the form

$$\begin{aligned}&[\omega'(x_1+0) - \omega'(t_1-0)] + \cdots + [\omega'(t_k+0) - \omega'(x_2-0)] \\ &\qquad = (x_1 - t_1)\omega''(\xi_1) + \cdots + (t_k - x_2)\omega''(\xi_k), \\ &\qquad\qquad (x_1 < \xi_1 < t_1, \cdots, t_k < \xi_k < x_2)\end{aligned}$$

and now conclude that it is negative. It follows that C lies under the curve $Y=\phi(x)$.

More precisely, if x_1 and x_2 are any two roots of the equation $m(x)=\phi(x)$ we may conclude that either the strict inequality $m(x)<\phi(x)$ holds for all $x_1<x<x_2$ or $m(x)\equiv\phi(x)$ in the whole interval. For if x' is a common point between x_1 and x_2, the development of $\omega(x)$ at x' must be of the form $\omega(x)=-c(x-x')^{2k}+\cdots$, $c>0$, $k\geqq 1$, whence $\omega''(x)=-2k(2k-1)c(x-x')^{2k-2}+\cdots$, and this is clearly incompatible with the inequality $\omega''(x)\geqq\omega(x)$.

The particular solution $\phi(x)$ which intersects $Y=m(x)$ at x_1 and x_2 is calculated from the equation

$$\begin{vmatrix} \phi(x) & e^{x} & e^{-x} \\ m(x_1) & e^{x_1} & e^{-x_1} \\ m(x_2) & e^{x_2} & e^{-x_2} \end{vmatrix} = 0.$$

Expressing the inequality $m(x)\leqq\phi(x)$ we get

THEOREM 2. *For* $x_1<x<x_2$

$$\begin{vmatrix} m(x) & e^{x} & e^{-x} \\ m(x_1) & e^{x_1} & e^{-x_1} \\ m(x_2) & e^{x_2} & e^{-x_2} \end{vmatrix} \geqq 0 \tag{4}$$

or

$$m(x) \leqq \frac{m(x_1)\sinh(x_2-x)+m(x_2)\sinh(x-x_1)}{\sinh(x_2-x_1)}.$$

If the sign of equality holds for one x *between* x_1 *and* x_2, *then it holds for all. Supposing* $m(x_1)$ *and* $m(x_2)>0$ *the equality holds if and only if* $|f(s)|>1$ *at all points between* x_1 *and* x_2.

The last assertion is an immediate consequence of the remark concerning the sign of equality in the condition (3).

The simplest functions for which the limits are actually attained are those of the form

$$f(s) = \exp(\alpha e^{s} + \beta e^{-s}),$$

where α and β are constants. If, for example, $\alpha>0$ and $\beta<0$ it should be noticed that $m(x)\equiv 0$ for $x<\frac{1}{2}\log(-\beta/\alpha)$, and that $m'(x)$ has a jump at the endpoint of this interval.

3. In order to study the possible shapes of the curve C: $Y=m(x)$ we fix an arbitrary point x_0 with $m(x_0)>0$. For convenience we suppose that $x_0=0$

and write $m_0=m(0)$, $m_0'=m'(0)$. The curves $Y=\phi(x)$ passing through the point $(0, m_0)$ are given by $\phi(x)=\alpha e^x+(m_0-\alpha)e^{-x}$, the slope at $x=0$ being $2\alpha-m_0$. Through every point with $x\neq 0$ passes one and only one curve of the field with $\alpha=\alpha(x)$. For $0<\alpha<m_0$ the curves tend towards infinity in both directions. The two curves $\alpha=0$ and $\alpha=m_0$ are exponential curves, and the curves $\alpha<0$ or $>m_0$ intersect the x-axis at $x=\frac{1}{2}\log[(\alpha-m_0)/\alpha]$.

The slope m_0' at $x=0$ corresponds to $\alpha=\frac{1}{2}(m_0+m_0')$. From Theorem 2 we infer that C does not cut any of the curves $\alpha>\frac{1}{2}(m_0+m_0')$ in the half-plane $x>0$, nor any of the curves $\alpha<\frac{1}{2}(m_0+m_0')$ in the half-plane $x<0$. Moreover, the values of $\alpha(x)$ are steadily non-decreasing as x increases. Hence $\alpha(x)$ must tend to definite limits as x tends to plus or minus infinity.

From these remarks we can easily deduce all the statements in

THEOREM 3. *For every C one of the following mutually exclusive statements is true*:

(a) $\eta_1=\lim_{x\to\infty}m(x)e^{-x}$ *and* $\eta_2=\lim_{x\to-\infty}m(x)e^{x}$ *exist and are positive*;

(b) $\lim_{x\to-\infty}m(x)=0$ *and* $\eta_1=\lim_{x\to\infty}m(x)e^{-x}>0$;

(c) $\lim_{x\to\infty}m(x)=0$ *and* $\eta_2=\lim_{x\to-\infty}m(x)e^{x}>0$.

If $|m_0'|<m_0$, *C is of type* (a); *if* $m_0'\geqq m_0$, *C is of type* (a) *or* (b); *and if* $m_0'\leqq -m_0$, *C is of type* (a) *or* (c).

The proof follows immediately from the fact that α is increasing. The normal situation in the cases (b) and (c) is of course that $m(x)$ vanishes identically from a certain point on.

If C is known to be of the type (b), then by Theorem 3 the inequality $m'(x)\geqq m(x)$ must hold for all x, for $x=0$ was only a representative for an arbitrary point. This differential inequality is equivalent with the fact that $m(x)e^{-x}$ is a non-decreasing function. Hence we obtain the more precise

THEOREM 4. *If the curve C is of type* (b) *the function $m(x)e^{-x}$ is non-decreasing, and for a curve of type* (c) *$m(x)e^{x}$ is non-increasing.*

All these results refer to the $m(x)$ of a single region Ω. The same proof can, however, be given also if $m(x)$ is formed with respect to any number of regions Ω, including the case considered by the brothers Nevanlinna, where $m(x)$ refers to all regions Ω.

4. We shall finally interpret our main result for the more familiar case of a function analytic in a half-plane.

THEOREM 5. *Let the function $f(z)$ be analytic in the closed right half-plane and suppose that $|f(z)|\leqq 1$ on the imaginary axis. If we write*

$$m(r)=\int_{-\pi/2}^{\pi/2}\overset{+}{\log}|f(re^{i\phi})|\cos\phi\,d\phi$$

the function $m(r)/r$ is non-decreasing and consequently tends to a positive limit as $r\to\infty$, unless $m(r)\equiv 0$.

The proof is obtained by taking $\log z$ as independent variable. There is only one alternative, since the curve C must obviously be of the type (b).

Remark. It is interesting to show that $m(r)\equiv\eta r$ implies $f(z)\equiv Ae^{2\eta z}$, where A is a constant of absolute value 1. In §4 we already proved that $Y=m(x)$ coincides with a curve $Y=\phi(x)$ if and only if $|f|>1$ for all interior points. Hence we must have $|f(x)|>1$ in the whole half-plane and the function can be continued to the left half-plane with $|f(z)|<1$. The entire function $f(z)$ has no zeros and is easily seen to be of the first order at most. Consequently it is of the form Ae^{cz}, where $|A|=1$ and c must be real and positive. The computation of $m(r)$ yields $c=2\eta$.

In the classical principle of Phragmén-Lindelöf the object of consideration is the maximum modulus $M(r)=\max_{|z|=r}|f(z)|$. Comparing with $m(r)$ we have $M(r)\geqq\frac{1}{2}m(r)$, whence we obtain

$$\liminf_{r\to\infty}\frac{\log M(r)}{r}>0.$$

Whether the limit of $(\log M(r))/r$ always exists we have not been able to decide.

The conditions of Phragmén-Lindelöf's principle are usually weakened by supposing that $f(z)$ is analytic only in the open half-plane and that $\limsup|f(z)|\leqq 1$ as z approaches a point on the imaginary axis. The most convenient way of extending Theorem 1, and hence all subsequent results, to this case would be first to consider a slightly smaller rectangle $x_1<x<x_2$, $|y|<\pi/2-\epsilon$ and to apply the theorem to the function $f(s)/\mu$, where μ is the upper bound of $|f|$ on the horizontal sides of the rectangle. A simple passage to the limit would then yield the desired result.

Part II

5. Phragmén-Lindelöf's principle is essentially a theorem on harmonic functions of two variables. It is natural to ask whether a similar theorem holds for harmonic functions of n variables. We shall show that our method can easily be generalized to this case.

Let the function $u(x_1,\cdots,x_n)$ be harmonic in some part of the n-dimensional space and consider one of the regions Ω in which u is positive. We shall suppose that Ω lies entirely in the half-space $x_1\geqq 0$. In addition we require that u shall be regular, and consequently $=0$, at all boundary points of Ω, except possibly at the origin and at infinity. The problem consists in studying the behavior of u in the region Ω.

Consider the intersection Δ_r or Ω with the sphere $x_1^2+x_2^2+\cdots+x_n^2=r^2$. In close analogy with the two-dimensional case we define

$$m(r) = \int_{\Delta_r} u \cos\phi \, d\omega,$$

where $\sin\phi = x_1/r$ and $d\omega$ is the central projection of the surface element on the sphere of radius 1.

A first differentiation is easy to carry out and yields

$$m'(r) = \int_{\Delta_r} \frac{\partial u}{\partial r} \cos\phi \, d\omega$$

since u vanishes on the boundary of Ω. The direct computation of the second derivative is rather intricate, so we prefer to make use of Green's formula which is well known to hold in n dimensions.

According to Green's formula we have

$$0 = \int \left(u \frac{\partial x_1}{\partial n} - x_1 \frac{\partial u}{\partial n} \right) d\sigma,$$

when the integral is extended over a closed surface. Take this surface to be the boundary of the part of Ω lying in the region $r_1 < r < r_2$ and let the normals be outer normals. Then Δ_{r_1} and Δ_{r_2} contribute together

$$\Big|_{r_1}^{r_2} r^{n-1}m(r) - r^n m'(r).$$

On the boundary of Ω we have $u=0$, $x \geqq 0$, and $\partial u/\partial n \leqq 0$. The corresponding part of the integral is thus seen to be non-positive, and we conclude that

$$r^{n-1}(m(r) - rm'(r))$$

is a non-increasing function of r.

In order to get a simple differential inequality we introduce $x = \log r$ as independent variable. We then get

$$\frac{d}{dx}\left[e^{(n-1)x} \left(m - \frac{dm}{dx} \right) \right] = e^{(n-1)x} \left((n-1)m - (n-2)\frac{dm}{dx} - \frac{d^2m}{dx^2} \right) \leqq 0$$

and finally

$$\frac{d^2m}{dx^2} + (n-2)\frac{dm}{dx} - (n-1)m \geqq 0.$$

This inequality holds whenever m'' exists. At irregular points one proves readily that m' is continuous or has a positive jump.

The easiest way of handling the differential inequality is to make the substitution $\mu = me^{(n/2-1)x}$. The inequality goes over into

$$\frac{d^2\mu}{dx^2} \geqq \left(\frac{n}{2}\right)^2 \mu,$$

and we are now in a position to apply the results already proved. In the first place we find that the curve C: $Y = \mu(x)$ is convex with respect to the family of curves $Y = \alpha e^{nx/2} + \beta e^{-nx/2}$. Expressing this result in terms of r we get

THEOREM 7. *In n dimensions the relation* (4) *is replaced by*

$$\begin{vmatrix} m(r)r^{n/2-1} & r^{n/2} & r^{-n/2} \\ m(r_1)r_1^{n/2-1} & r_1^{n/2} & r_1^{-n/2} \\ m(r_2)r_2^{n/2-1} & r_2^{n/2} & r_2^{-n/2} \end{vmatrix} \geqq 0.$$

We note that the sign of equality holds for harmonic functions of the form $\alpha x_1 + \beta x_1/r^n$.

Theorem 3 is immediately carried over if we replace $m(x)$ by $\mu(x)$ and the exponentials $e^{\pm x}$ by $e^{\pm nx/2}$. The numbers η_1 and η_2 of the theorem are thus defined by $\eta_1 = \lim_{r\to\infty} m(r)/r$ and $\eta_2 = \lim_{r\to 0} m(r)r^{n-1}$.

We conclude the paper by a complete statement of the assertions corresponding to Theorems 4–5 of the first section.

THEOREM 8. *Let $u(x_1, \cdots, x_n)$ be harmonic and positive in a region Ω contained in the half-space $x_1 \geqq 0$. Suppose further that u is regular and $=0$ at all boundary points of Ω, except possibly at the origin and at infinity. The expression*

$$m(r) = \int_{\Delta_r} u \cos\phi \, d\omega$$

has the following properties:

The limits

$$\eta_1 = \lim_{r\to\infty} \frac{m(r)}{r} \quad \text{and} \quad \eta_2 = \lim_{r\to 0} m(r)r^{n-1}$$

exist and one at least is positive.

If $\eta_1 = 0$ the function $m(r)/r$ is non-decreasing, and $\eta_2 = 0$ implies that $m(r)r^{n-1}$ is non-increasing.

We remark that the maximum modulus $M(r)$ is greater than $m(r)$ multiplied by a certain constant. Hence we conclude that one at least of the relations $\lim_{r\to\infty} M(r)/r > 0$ and $\lim_{r\to 0} M(r)r^{n-1} > 0$ is true. The extremal functions are x_1 and x_1/r^n. In two dimensions one of these is obtained from the other by an inversion, but in higher dimensions they are essentially different.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

EXTENSIONS OF THE FOUR-VERTEX THEOREM*

BY

W. C. GRAUSTEIN

1. **Introduction.** The "Vierscheitelsatz" states that an oval has at least four vertices, that is, that the curvature of an oval has at least four relative extrema.† It is the purpose of this paper to investigate the extent to which the theorem can be generalized.

In order that the theorem make geometric sense, it is essential that the curve under consideration be closed and have continuous curvature. Beyond these assumptions the only restriction placed on the curve is that it be regular. Initially, to be sure, it is assumed that the curve contains no rectilinear segments, but this restriction is eventually lifted.

By the angular measure of a regular curve shall be meant the algebraic angle through which the directed tangent turns when the curve is completely traced. The angular measure of a closed regular curve is $2n\pi$, where n is zero or an integer which can be made positive by a suitable choice of the sense in which the curve is traced. Closed regular curves may thus be divided into classes K_n $(n=0, 1, 2, \cdots)$, where the class K_n consists of all the curves of angular measure $2n\pi$.

A closed regular curve is an oval if (a) it has no points of inflection and either (b_1) it is simple or (b_2) it is of class K_1. Of the latter conditions we shall choose the condition (b_2). This condition is, in fact, weaker than the alternative condition (b_1). In other words, the simple closed regular curves constitute a subclass of the curves K_1.‡

It turns out that the four-vertex theorem is true for every simple closed regular curve except, of course, a circle. Fundamental in the establishment of this fact is the analysis of a certain kind of arc which, on account of the similarity of its shape to the curved portion of the Greek capital omega, is called an arc of type Ω. The most significant property, for the purpose in view, of an arc of this type is that, when it is traced so that its curvature is non-negative,

* Presented to the Society, September 3, 1936; received by the editors June 16, 1936.

† The theorem seems to have been first published by Mukhopadhyaya, *New methods in the geometry of a plane arc*, Bulletin of the Calcutta Mathematical Society, vol. 1 (1909), pp. 31–37. For other proofs, see, for example, Kneser, *Bemerkungen über die Anzahl der Extreme der Krümmung auf geschlossenen Kurven*, Heinrich Weber Festschrift (1912), pp. 170–180; Blaschke, *Die Minimalzahl der Scheitel einer geschlossenen konvexen Kurve*, Rendiconti di Palermo, vol. 36 (1913), pp. 220–222; Bieberbach, *Differentialgeometrie*, pp. 23–27.

‡ See, e.g., H. Hopf, *Über die Drehung der Tangenten und Sehnen ebener Kurven*, Compositio Mathematica, vol. 2 (1935), pp. 50–62.

the curvature has always at least one minimum interior to the arc (Theorem Ω, §4). By means of this property the four-vertex theorem is established for ovals which contain rectilinear segments (§6) as well as for ordinary ovals (§5). To establish the theorem for every simple closed regular curve it is found necessary to bring in, also, the theory of directed lines of support (§§7, 8).

Examples are given in §11 to show that the four-vertex theorem is not true for all the curves of any given class, K_n, of closed curves. However, there exists in each one of these classes a subclass of curves for which the theorem is true. When $n \geqq 2$, this subclass consists of all the curves which contain an arc of type Ω and possess points of inflection. If $n=0$ or $n=1$, the subclass consists of all the curves containing an arc of type Ω; the prescription of points of inflection is unnecessary in either of these cases. It is to be noted that the subclass of the class K_1 thus defined contains, in addition to the simple closed curves, many types of curves with double points.

In a recent paper* the four-vertex theorem for an oval was given a new form which is more discriminating and also more in keeping with the theorem as a result in the large. The theorem in its new form states that an oval has at least four *primary* vertices. By a primary vertex is meant an extremum of the curvature which is greater than or less than the average curvature according as it is a maximum or a minimum. The present paper, whose methods are largely qualitative rather than quantitative, deals with the theorem in its original form. Extension of the theorem in its new form is left an open question.

2. **Regular plane curves.** A plane curve admitting a parametric representation in terms of real single-valued functions of a real variable which are of class C'' in a certain interval and whose first derivatives never vanish simultaneously in this interval is known as a regular plane curve of class C'', and the parameter in question is called a regular parameter. For such a curve the arc s is a regular parameter and hence the curve has regular parametric equations of the form $x=x(s)$, $y=y(s)$.

If $\phi=\phi(s)$ is the single-valued function of s of class C' which measures the directed angle from the positive axis of x to the directed tangent to the curve at the point $P:s=s$, the curvature at P is given by $1/R=d\phi/ds$.

The curvature at P is positive or negative according as the curve in an immediate neighborhood of P lies to the left or to the right of the directed tangent at P, is zero if P is a point of inflection, and changes sign when the direction in which the arc is measured is reversed.

* Graustein, *A new form of the four-vertex theorem*, Monatshefte für Mathematik und Physik. Wirtinger Festband (1936), pp. 381–384.

From the definition of ϕ, it follows that $dx/ds = \cos\phi$, $dy/ds = \sin\phi$. Hence, we obtain, as regular parametric equations of the curve,

$$(1) \qquad x = \int_0^s \cos\phi\, ds, \qquad y = \int_0^s \sin\phi\, ds,$$

where

$$(2) \qquad \phi = \int_0^s \frac{ds}{R},$$

provided merely that the point $P_0 : s = 0$ is taken as the origin of coordinates and the directed tangent at P_0 as the positive axis of x.

The angle ϕ is, by definition, the directed arc of the tangent indicatrix, $\xi = \cos\phi$, $\eta = \sin\phi$, of the given curve. This arc is measured positively in the counterclockwise direction on the circle which bears the tangent indicatrix and negatively in the opposite direction.

By the *angular measure* of the arc P_1P_2 of the given curve, where P_1 and P_2 are the points $s = s_1$ and $s = s_2$ respectively, we shall mean the quantity $\phi_2 - \phi_1$, where ϕ_i is the value of ϕ for $s = s_i$, $i = 1, 2$. This angular measure may be thought of geometrically either as the algebraic angle through which the tangent to the curve at P turns when P traces the arc P_1P_2 from P_1 to P_2 or as the algebraic arc length of corresponding arc of the tangent indicatrix, when correspondingly traced. It may be positive, negative, or zero. The angular measure of a circle is, for example, $\pm 2\pi$, depending on the direction in which the circle is traced, and the angular measure of a lemniscate is zero.

3. **Arcs of non-negative curvature.** Consider a regular curvilinear arc of class C'' which contains no rectilinear segments and whose curvature is of one sign or zero. Trace the arc in the sense which makes the curvature non-negative, and denote the initial and terminal points of the arc so traced by A and B respectively. Measure the directed arc length from A and choose coordinate axes so that the complete arc, AB, is represented by equations (1), $0 \leqq s \leqq l$, where l is the length of the arc.

It will be convenient to use, instead of (1), the parametric representation of the arc AB in terms of the angle ϕ. Since $d\phi/ds \geqq 0$, $0 \leqq s \leqq l$, and the arc AB contains no rectilinear segments, $\phi = \phi(s)$ is a non-decreasing function of ϕ and $s = s(\phi)$ is a single-valued function of ϕ, defined over the interval $0 \leqq \phi \leqq \lambda$, where λ is the angular measure of the arc AB. Thus, the desired parametric representation is

$$(3) \qquad x = \int_0^\phi R(\phi) \cos\phi\, d\phi, \qquad y = \int_0^\phi R(\phi) \sin\phi\, d\phi, \qquad 0 \leqq \phi \leqq \lambda.$$

At the points of the interval $0 \leqq \phi \leqq \lambda$ at which $1/R=0$, the integrals in (3) are improper but convergent. Except at these points, $R=R(\phi)$ is a continuous function of ϕ and ϕ is a regular parameter for the curve.

LEMMA 1. *Let AB be an arc of non-negative curvature with no rectilinear segments, represented by equations of the form* (3), *and let $P_0=A, P_1, P_2, P_3, \cdots$, be the points on it for which $\phi=0$, $\phi=\pi/2$, $\phi=\pi$, $\phi=3\pi/2, \cdots$, respectively. Then the arc P_0P_1 is rising to the right, concave upward; the arc P_1P_2 is rising to the left, concave downward; the arc P_2P_3 is falling to the left, concave downward; the arc P_3P_4 is falling to the right, concave upward; and so forth.**

From the relations (3) we have $dx/d\phi=R\cos\phi$, $dy/d\phi=R\sin\phi$, and $d^2y/dx^2=\sec^3\phi/R$, and hence the lemma follows.

LEMMA 2. *The arc AB of Lemma* 1, *if it does not cut itself and never gets below the tangent L at A, has the character of an inwinding spiral. The only point on it, other than A, at which the tangent can coincide with L is the point $P_4:\phi=2\pi$, and P_4 is then to the left of A.*

Let the tangents at the points $P_0=A, P_1, P_2, \cdots$, defined in Lemma 1 be $T_0=L, T_1, T_2, \cdots$, respectively. By Lemma 1, T_2 lies above T_0 and T_4 lies below T_2. If T_4 were also below T_0, the arc P_2P_4 would cut T_0 and the hypothesis that AB never gets below T_0 would be contradicted. Thus, T_4 lies below T_2 and lies above or is coincident with T_0. Therefore, T_4 cuts the arc P_0P_2 in a unique point C, other than P_2.

The point P_4 lies to the left of C on T_4 and hence to the left of P_0 if $T_4=T_0$ and $C=P_0$. For, if P_4 were to the right of C or coincided with C, P_4 would be outside the open region S_1 bounded by the arc P_0P_2, the lines T_0, T_2 and any line parallel to T_3 which is to the left of both T_3 and A, whereas the point $P:\phi=\pi+\epsilon$, $\epsilon>0$, is, for ϵ sufficiently small, in S_1; but, by Lemma 1, the arc P_2P_4 cannot leave S_1 except over the arc P_0P_2 and the arc AB would then have cut itself.

Consider, next, the open region S_2 bounded by the arc CP_2P_4 and the line segment P_4C. The arc P_4P_6 is entirely in S_2 since, by Lemma 1, it is above the line $T_4=P_4C$ and cannot intersect the arc CP_2P_4. In particular, then, T_5 is to the left of T_1, and T_6 is below T_2.

The argument now begins to repeat itself. Since P_6 is in S_2, T_6 cuts the arc P_2P_4 in a point D to the left of P_6, and the arc P_6P_8 lies in the open region S_3 bounded by the arc DP_4P_6 and the line segment P_6D, and so forth. Thus, the lemma is fully established.

* In using here, the terms "right," "left," "upward," "downward" and similar terms, we assume that the coordinate axes to which the arc AB is referred are so visualized that they appear horizontal and vertical respectively, with the x-axis directed to the right and the y-axis directed upward.

4. **Arcs of type Ω.** We shall say that an arc AB is of type Ω if (a) its curvature (when it is traced from A to B) is non-negative and it contains no rectilinear segments; (b) the tangents at A and B are identical; (c) it lies on one side of the common tangent, L, at A and B; and (d) it does not cut itself except that B may coincide with A.

As in §3, we shall assume that the arc AB is referred to A as origin of coordinates and to the directed tangent at A as axis of x, and we shall think of it as represented by equations of the form (3).

LEMMA 3. *The angular measure of an arc AB of type Ω is 2π and the terminal point B lies to the left of the initial point A, if it is not coincident with it.*

These facts follow directly from Lemma 2.

LEMMA 4. *The curvature of an arc AB of type Ω cannot be monotonic, unless it is constant.*

The points A and B have respectively the coordinates $y_A = 0$ and

$$y_B = \int_0^\pi R \sin \phi \, d\phi + \int_\pi^{2\pi} R \sin \phi \, d\phi.$$

It is evident that, if R is not constant, y_B would be positive if R were non-increasing, and negative if R were non-decreasing. But, by property (b) of an arc of type Ω, $y_B = y_A = 0$. Thus, the lemma is proved.

DEFINITION. An arc AB of angular measure 2π whose curvature (when the arc is traced from A to B) is non-negative is said to overlap itself with respect to a point K on it if, when K is taken as the origin of new coordinates $(\bar{x}, \bar{y})$ and the directed tangent at K is chosen as the $\bar{x}$-axis, the point B is horizontally to the left of the point A with respect to the new axes: $\bar{x}_A - \bar{x}_B > 0$.

LEMMA 5. *An arc AB of type Ω fails to overlap itself with respect to every point on it for which $\pi/2 \leqq \phi \leqq 3\pi/2$. It does overlap itself with respect to every point on it for which $0 \leqq \phi < \pi/2$ or $3\pi/2 < \phi \leqq 2\pi$, provided merely that B does not coincide with A.*

If $B = A$, there is evidently no overlap with respect to any point of the arc. If $B \neq A$, the results stated follow directly from the fact that the tangents at B and A are identical and B is to the left of A (Lemma 3).

LEMMA 6. *If the curvature of an arc AB of type Ω has just one extremum interior to the arc AB, this extremum must be a minimum.*

By an extremum of the curvature shall be meant a point of the arc AB, or of the corresponding interval $0 \leqq \phi \leqq 2\pi$, at which the curvature has a relative extremum, or a segment of the arc, or the interval, for which the curva-

ture has a constant relatively extreme value with respect to neighboring segments on either side of it. In the latter case, the extremum shall be said to be interior to the arc AB, or the interval $0 \leqq \phi \leqq 2\pi$, only if the segment in question contains neither end point of the arc, or the interval.

We shall establish the desired conclusion by showing that the opposite one leads to a contradiction. If the single extremum of $1/R$ is a maximum, $1/R$ can vanish only at A or at B. Hence, R is continuous interior to the arc AB, has a single extremum, a minimum, interior to the arc AB, and may become infinite at A or at B.

The curve which consists of the graph of the function $R = R(\phi)$, $0 \leqq \phi \leqq 2\pi$, and the portions of the lines $\phi = 0$ and $\phi = 2\pi$ which lie respectively above the points $\phi = 0$ and $\phi = 2\pi$ of this graph when these points exist, is a continuous curve and hence possesses a horizontal chord of length π. It follows from the conditions on $R(\phi)$ that, if c is the height of this chord above the y-axis, $R(\phi) \leqq c$ for the portion of the graph having the same projection on the ϕ-axis as the chord and $R(\phi) \geqq c$ for the remainder of the graph. Hence, if γ is the abscissa of the midpoint of the chord,

$$I = \int_0^{2\pi} R(\phi) \cos (\phi - \gamma) d\phi < 0.$$

When we set $\phi - \gamma = \bar{\phi}$, the integral I becomes

$$I = \int_0^{2\pi - \gamma} \bar{R}(\bar{\phi}) \cos \bar{\phi}\, d\bar{\phi} - \int_0^{-\gamma} \bar{R}(\bar{\phi}) \cos \bar{\phi}\, d\bar{\phi} = \bar{x}_B - \bar{x}_A,$$

where $\bar{x}_A$ and $\bar{x}_B$ are respectively new abscissas of the points A and B, referred to the point $K\colon \phi = \gamma$ on the arc AB as origin and the directed tangent at K as axis of $\bar{x}$. Hence, $\bar{x}_A - \bar{x}_B > 0$. By construction, $\pi/2 \leqq \gamma \leqq 3\pi/2$. Thus, the arc AB overlaps itself with respect to a point on it for which $\pi/2 \leqq \phi \leqq 3\pi/2$. But, herewith we have the desired contradiction to Lemma 5.

THEOREM Ω. *The (non-negative) curvature of an arc of type Ω has at least one minimum interior to the arc or is constant throughout the arc.*

Assuming that the curvature is not constant, we conclude that it has extrema in the closed interval $0 \leqq \phi \leqq 2\pi$, since it is continuous in this interval. Furthermore, it is not monotonic, by Lemma 4, and hence has at least one extremum in the open interval. If it has just one, this must be a minimum, by Lemma 6, and if it has more than one, at least one must be a minimum in any case.

The arc (3) for which

$$R = a + \sin (\phi/2), \qquad a > 1, \qquad 0 \leqq \phi \leqq 2\pi,$$

is readily shown to be an arc of type Ω with a single extremum, a minimum, interior to it.

5. **Ovals.** By an oval we shall mean here a closed regular curve of class C'' with no rectilinear segments whose curvature is always of one sign or zero and whose tangent indicatrix is a circle traced just once. We shall assume that the oval is so traced that the curvature is non-negative. The angular measure is then 2π.

LEMMA 7. *An oval is an arc of type Ω with respect to every point on it.*

If A is a point of an oval, the oval may be considered as an arc AB of non-negative curvature, where $B=A$. It will follow that this arc is of type Ω if it can be shown that it lies to the left of the tangent at A and does not cut itself. But these properties follow without difficulty by application of Lemma 1.

By a vertex of a closed curve we shall mean an extremum of the curvature, as defined in §4.

THEOREM 1. *An oval, not a circle, has at least four vertices.*

The curvature of the oval has in any case two extrema, a maximum and a minimum. Let A be the point of the oval, or a point of the segment thereof, for which the curvature has a minimum. By Lemma 7, the arc AB, where $B=A$, is of type Ω, and hence, by Theorem Ω, its curvature has at least one minimum interior to it. Thus, the curvature of the oval has at least two minima and therefore at least two maxima.

6. **Flattened ovals.** A closed regular curve of class C'' which contains one or more rectilinear segments and, when properly traced, has non-negative curvature and angular measure 2π, we shall call a flattened oval.

It may be readily proved by use of an extension of Lemma 1 covering the case of an arc of non-negative curvature with rectilinear segments that a flattened oval lies to the left of every directed tangent and does not cut itself. Hence, we conclude the following proposition.

LEMMA 8. *A flattened oval with only one rectilinear segment may be thought of as consisting of an arc AB of type Ω and the line segment BA.*

The curvature of a flattened oval of this type has a minimum on the line segment BA and, by Theorem Ω, at least one minimum interior to the arc AB; and that of a flattened oval with more than one rectilinear segment has at least as many minima as there are rectilinear segments. Thus, we pass to the following result.

THEOREM 2. *A flattened oval has at least four vertices.*

7. **Lines of support.** Let C be a closed regular curve of class C'' without component line segments.

LEMMA 9. *If C, or an open arc of C, lies wholly on one side of a line L except for one or more points on L, then L is a non-inflectional tangent to C at each of these points.*

By hypothesis, C has neither "corners" nor cusps. Hence, in passing through a point P, C crosses every line through P, including the tangent at P, unless this tangent is non-inflectional.

DEFINITION. A directed line L shall be called a line of support of the curve C if C is tangent to L and lies wholly to the left of L except for its one or more points of contact with L.

LEMMA 10. *The curve C has a unique line of support in every oriented direction.*

A pencil of parallel and similarly directed lines admits a unique dichotomy such that all points of C are to the left of every line of the first class and at least one point of C lies on, or to the right of, an arbitrarily chosen but fixed line of the second class. Clearly, there is no "last" line in the first class. There is then a "first" line, L, in the second class. Evidently, C lies wholly to the left of L except for one or more points on it and, therefore, by Lemma 9, C is tangent to L at each of these points. Thus, L is the unique line of support in the given oriented direction.

DEFINITION. A line of support of the curve C shall be called a simple, or a multiple, line of support, according as it is tangent to C in one point or more than one point.

LEMMA 11. *If the directed tangent T_0 to the directed curve C at a point P_0 is a simple line of support, the directed tangent to C at every point in a certain neighborhood of P_0 is a simple line of support. In this neighborhood of P_0 there exists a parametric representation of C in terms of the directed angle from T_0 to an arbitrary line of support and this parametric representation is identical with a representation in terms of the arc of the tangent indicatrix.*

We refer C to P_0 as origin and T_0 as x-axis, and measure the arc s from P_0. By hypothesis, the curvature of C at P_0 is positive or zero, and, if zero, does not change sign at P_0. Hence $\epsilon>0$ exists so that at all points of the arc $P_{-\epsilon}P_0P_\epsilon$, that is, at every point P_s of C for which $-\epsilon\leqq s\leqq\epsilon$, the curvature is non-negative. It follows, by Lemma 1, that, if T_s is the directed tangent at the arbitrary point P_s of this arc, the arc lies to the left of T_s except for the point P_s, provided that ϵ is so chosen that the angular measure of the arc is less than π.

The directed distance $D(s, \bar{s})$ from T_s to a variable point on the complementary arc $P_\epsilon P_{-\epsilon}$, that is, to an arbitrary point $P_{\bar{s}}$ for which $\epsilon \leqq \bar{s} \leqq l-\epsilon$, where l is the length of C, is a continuous function of s and $\bar{s}$. Moreover, it follows from the fact that T_0 is a simple line of support that $D(0, \bar{s})$, the directed distance from T_0 to $P_{\bar{s}}$, is bounded away from zero for $\epsilon \leqq \bar{s} \leqq l-\epsilon$. Hence, a positive constant $\delta \leqq \epsilon$ exists so that for every T_s for which $-\delta < s < \delta$, $D(s, \bar{s})$ fails to vanish for $\epsilon \leqq \bar{s} \leqq l-\epsilon$. Consequently, the tangent T_s at every point P_s for which $-\delta < s < \delta$ is a simple line of support, and the first part of the lemma is proved.

The arc $P_{-\delta}P_0P_\delta$, since its curvature is non-negative, admits a representation in terms of the directed angle ϕ from the positive x-axis to an arbitrary directed tangent; see §2. But the directed tangents to the arc are all simple lines of support of C and constitute precisely the lines of support of C in the oriented directions of these tangents. Hence, the angle ϕ can be thought of as the angle from the positive axis of x to an arbitrary one of these lines of support, and the second part of the lemma is established.

8. **Simple closed curves.** We shall mean by a simple closed curve a closed regular curve of class C'' which contains no rectilinear segments and no double points.

LEMMA 12. *A simple closed curve all of whose lines of support are simple is an oval.*

Lemma 11 implies in this case that the lines of support of the given curve C envelope a closed continuous arc of C which when properly traced has non-negative curvature and angular measure 2π; in other words, an oval. But C is a connected curve without double points and consists, therefore, only of this oval.

LEMMA 13. *A simple closed curve, not an oval, has at least one multiple line of support.*

This follows directly from Lemma 12.

LEMMA 14. *The directions of a directed simple closed curve at two points of contact with a multiple line of support are the same and the number of inflections in an arc of the curve joining the two points, unless infinite, is even or zero.*

Consider the open region S bounded by the multiple line of support, L, and one of the two arcs of the given curve C joining the two given points of contact of C with L. If the directions of C at the two given points were opposite, there would exist two points neighboring respectively to the given points and belonging to the second arc of C joining the given points, one of which is inside S while the other is outside S, and hence the second arc of C

would have to cross either the first arc or L. In either case an hypothesis would be contradicted. Thus, C has the same direction at the two given points.

Let C now be directed so that the common direction at the two points is that of L. Then C lies to the left of its directed tangents at the two points. Moreover, since L is a line of support, neither of the points is an inflection. Therefore, in sufficiently restricted neighborhoods of the two points, the curvature of C is non-negative and hence the number of inflections in each arc of C joining the two points, if not infinite, is even or zero.

LEMMA 15. *If an arc of a simple closed curve which joins two points of contact of the curve with a multiple line of support contains no points of inflection, the arc, when suitably traced, is of type* Ω.

Let the given curve C be directed as in the paragraph preceding the lemma, and denote by A and B respectively the initial and terminal points of the arc of C in question. Then the arc AB has non-negative curvature, and obviously has the remaining properties required of an arc of type Ω.

LEMMA 16. *A simple closed curve with no points of inflection is an oval.*

It suffices to show that all the lines of support of the given curve C are simple, for Lemma 12 will then apply. Suppose there were a multiple line of support, L, tangent to C at the distinct points A and B. Then, if C were properly traced, both of the arcs AB and BA would be of type Ω, by Lemma 15. Consequently, if L were thought of as horizontal, each of the points A and B would be to the left of the other on L, by Lemma 3. This is absurd, and accordingly C has only simple lines of support.

9. **The four-vertex theorem for simple closed curves.** We prove the following theorem:

THEOREM 3. *A simple closed curve, not a circle, has at least four vertices.*

A simple closed curve C obviously has no inflections, an even number, or infinitely many. Inasmuch as there is at least one vertex between each two points of inflection, the theorem is true, though trivial, if there are more than two points of inflection.

If there are no points of inflection, the curve C is an oval, by Lemma 16,* and hence has, according to Theorem 1, at least four vertices, unless it is a circle.

There remains the case in which there are just two points of inflection.

* It seems preferable to employ here the simple Lemma 16 rather than the more powerful proposition that every simple closed curve, properly traced, has angular measure 2π (see, e.g., Hopf, loc. cit.), for thereby the proof of Theorem 3 is effected without appeal to this proposition.

The curve C cannot, then, be an oval, and hence has, by Lemma 13, at least one multiple line of support, L. Let A and B be two points of contact of C with L, and consider the arcs AB and BA in which they divide C. According to Lemma 14, one of these arcs must contain both points of inflection of C. The other arc contains, then, no point of inflection and is therefore, by Lemma 15, of type Ω. Let this be the arc AB, traced from A to B, and denote the complementary arc now by BI_1I_2A, where I_1 and I_2 are the two points of inflection.

Since we are now tracing C so that the direction at A is that of the directed line L, the curvature of the arc I_2ABI_1 is non-negative, whereas that of the arc I_1I_2 is non-positive.

Inasmuch as the arc AB is of type Ω, it follows, from Theorem Ω, that $1/R$ has a minimum interior to the arc AB and hence interior to the arc I_2ABI_1. But $1/R$ surely has a minimum interior to the arc I_1I_2. Thus, $1/R$ has at least two minima and so must have at least two maxima. Consequently, C has at least four vertices.

10. **The four-vertex theorem for flattened simple closed curves.** A closed regular curve of class C'' without double points which contains one or more rectilinear segments we shall call a flattened simple closed curve.

THEOREM 4. *A flattened simple closed curve has at least four vertices.*

The proof of Theorem 3 rests ultimately on Lemmas 1–6, 9–16 and Theorem Ω. Theorem 4 will be established if it can be shown that these propositions remain valid in all essentials when rectilinear segments are admitted as component parts of the curves and arcs, including arcs of type Ω, which are discussed in them. This is indeed the case, as inspection, in view of the following remarks on the more pertinent points at issue, will show.

Though the given curve or arc does not admit, along the component rectilinear segments, a parametric representation in terms of the arc ϕ of the tangent indicatrix, the angle ϕ may still be thought of as a parameter for the curve or arc, provided it is agreed that to certain values of ϕ shall correspond, not points, but segments of straight lines. In keeping with this agreement, the meaning of the term "point" must be broadened, on occasion, to include rectilinear segments. In particular, by a "point of inflection" may be meant either a point at which the curvature changes sign or a rectilinear segment along which the curvature changes sign with respect to neighboring arcs, one on each side of the segment. Again, a line of support will be simple if it is tangent to the given curve either in just one point or along just one rectilinear segment. In this connection, it should be noted that the parametric representation mentioned in Lemma 11 may fail in the sense above described.

However, the only proposition dependent on Lemma 11, namely, Lemma 12, is still valid. Of course, this proposition, and also Lemma 16, has to do with a *flattened* oval as well as with a flattened simple closed curve.

Inasmuch as the parametric representation (3) is employed to establish Lemmas 4 and 6, the possibility of the extension of these lemmas must be particularly scrutinized. Both lemmas deal with an arc AB of type Ω. There can be no rectilinear segment *interior* to AB by the hypothesis of Lemma 4, and, if there were one in the case of Lemma 6, this lemma would be obvious. It may, therefore, be assumed in both cases that a rectilinear segment occurs only at an end of the arc AB. When such segments are suppressed, the proofs of the original lemmas are valid and guarantee the desired extensions.

11. **Further extensions of the theorem.** In this paragraph we shall understand by a closed curve any closed regular curve of class C'', with or without rectilinear segments.

Curves of class K_1. We have shown that every simple closed curve, other than a circle, has at least four vertices. As is well known,* the simple closed curves form a subclass of the class K_1, consisting of all closed curves with angular measure 2π.†

It is not true that every curve of the class K_1 has at least four vertices, as will be evident shortly from an example. However, there exists a subclass of K_1 which includes, besides all simple closed curves, many types of curves with double points, for which the four-vertex theorem holds, namely, the subclass of curves which contain arcs of type Ω with or without rectilinear segments.

THEOREM 5a. *A closed curve of angular measure 2π which contains an arc of type Ω has at least four vertices or is a circle.*

Since the angular measure of an arc of type Ω, when properly traced, is 2π, the given curve is either an oval or has inflections. In the former case Theorems 1 and 2 apply, and in the latter, the general argument in the proof of Theorem 3 in the case of inflections is valid.

For the purpose of giving examples of curves of class K_1 which have just two vertices, we need conditions of closure for a curve when it is given by its intrinsic equation.

It is known that, if $1/R$ is a real, single-valued, continuous function of s in the interval $-\infty < s < \infty$, and

$$\phi = \phi(s) = \int_0^s \frac{ds}{R}, \tag{4}$$

* See, e.g., H. Hopf, loc. cit.

† Here, and later, we assume that every closed curve is so traced that its angular measure is non-negative.

then the equations

$$x = \int_0^s \cos\phi\, ds, \qquad y = \int_0^s \sin\phi\, ds \tag{5}$$

represent a regular plane curve of class C'' for which s is the measure of the arc and $1/R$ is the curvature. This curve is closed and of length l if the functions $x(s)$ and $y(s)$ in (5) are periodic with l as their smallest positive common period.

LEMMA 17. *A necessary and sufficient condition that the curve* (5) *be closed and of length l is that* $1/R$ *be a periodic function of s with the period l such that*

$$\phi(l) = 2n\pi, \qquad n \text{ an integer or zero}, \tag{6}$$

$$\int_0^l \cos\phi\, ds = 0, \qquad \int_0^l \sin\phi\, ds = 0, \tag{7}$$

and that l be the smallest positive period with these properties.

Suppose that the curve is closed and l is its length. Then, $x(l)=y(l)=0$ and equations (7) hold. Furthermore, differentiation of the identities

$$x(s+l) \equiv x(s), \qquad y(s+l) \equiv y(s) \tag{8}$$

leads to the relation,

$$\phi(s+l) \equiv \phi(s) + 2n\pi, \tag{9}$$

whence it follows that $\phi(l)=2n\pi$ and, on differentiation, that $1/R$ is periodic of period l.

Conversely, the assumption that $1/R$ is periodic of period l such that (6) holds yields (9) and, by means of (9) and (7), the identities (8) are readily established, and the curve is closed. Moreover, since l is the smallest positive period of $1/R$ for which (6) and (7) are valid, l is the length of the curve.

It is evident that $\phi(l)=2n\pi$ is the angular measure of the curve. In other words, the curve is of class K_n.

Consider, now, the curve with the intrinsic equation

$$\frac{1}{R} = a\cos s + 1, \qquad a > 0. \tag{10}$$

Inasmuch as $1/R$ is periodic with period 2π and has just two extrema in a period interval, and since

$$\phi(s) = a\sin s + s,$$

so that $\phi(2\pi)=2\pi$, this curve is a closed curve of class K_1 with just two vertices provided merely that a be so chosen that

$$I_1 = \int_0^{2\pi} \cos (a \sin s + s)ds = 0, \qquad I_2 = \int_0^{2\pi} \sin (a \sin s + s)ds = 0.$$

It is found that $I_2 \equiv 0$ and that $I_1 \equiv -4F(a)$, where

$$F(a) = \int_0^{\pi/2} \sin (a \sin s) \sin s\, ds.$$

When $\sin s$ is replaced by t and the integral is integrated by parts, it turns out that $F(a) = af(a)$, where

$$f(a) = \int_0^1 (1 - t^2)^{1/2} \cos at\, dt.$$

Hence equation (10) represents a closed curve of the desired type for every value of a for which $f(a) = 0$.

It is readily shown that the function $f(a)$ satisfies the differential equation

$$f''(a) + \frac{3}{a} f'(a) + f(a) = 0$$

and that $|f'(a)| \leqq 1/3$. Hence $|f''(a) + f(a)| \leqq 1/a$, and it follows that $f(a)$ is an oscillating function, with infinitely many zeros.

The smallest zero is approximately $(11/9)\pi$; the corresponding closed curve (10) has the shape of a figure eight with a loop in one lobe, interior to the lobe. The next zero is about $(41/18)\pi$ and the corresponding curve looks like a figure eight with two loops, one within the other, in the one lobe and a single loop in the other lobe. The curve corresponding to the nth zero has n loops in the one lobe, and $n-1$ in the other. All the curves are *analytic* and are symmetric in the y-axis.

Since a curve of class K_1 with no inflections is an oval and one with more than two inflections surely has at least four vertices, a necessary condition that a curve of class K_1 have just two vertices is that it have just two points of inflection, with the two vertices separated by them. It would appear, then, that every curve of class K_1 with just two vertices has the general character of one of the infinite set of the curves just described.

Curves of class K_0. The counterpart of Theorem 5a is true also in this case.

THEOREM 5b. *A closed curve of angular measure zero which contains an arc of type Ω has at least four vertices.*

For, since an arc of type Ω, properly traced, has angular measure 2π, the given curve has inflections and therefore the general argument in the proof of Theorem 3 applies.

A lemniscate is a simple example of a curve of class K_0 which has just two vertices.

Curves of class K_n, $n \geqq 2$. The analog of Theorem 5b (or 5a) is not true for a curve of angular measure greater than 2π, inasmuch as in this case it does not follow from the hypothesis of the theorem that the curve has inflections, and there actually exist curves without inflections which have only two vertices and yet contain an arc of type Ω. A limaçon with a double point is, for example, a curve of class K_2 with these properties.

It is, however, true, as is readily verified, that the following variation of Theorem 5b (or 5a) holds.

THEOREM 5c. *A closed curve of angular measure greater than 2π which has inflections and contains an arc of type Ω has at least four vertices.*

It is not difficult to give an example of a closed curve of positive curvature belonging to any prescribed class $K_n (n \geqq 2)$ and having just two vertices. For this purpose we note that, since we are assuming $1/R$ positive, equations (5) may be replaced by the equations

$$x = \int_0^\phi R \cos \phi \, d\phi, \qquad y = \int_0^\phi R \sin \phi \, d\phi, \qquad -\infty < \phi < \infty, \tag{11}$$

and that Lemma 17 then becomes:

LEMMA 18. *A necessary and sufficient condition that the curve of positive curvature* (11) *be a closed curve of class K_n is that n be the smallest positive integer such that $R = R(\phi)$ is periodic of period $2n\pi$ and*

$$\int_0^{2n\pi} R \cos \phi \, d\phi = 0, \qquad \int_0^{2n\pi} R \sin \phi \, d\phi = 0. \tag{12}$$

By means of this lemma it is readily verified that the curve with the intrinsic equation

$$R = a + \cos \frac{\phi}{n}, \qquad a > 1,$$

where n is an integer greater than unity, is a closed curve of class K_n. That the curve has just two vertices is obvious. In particular, if $n = 2$, the curve has the same general character as the limaçon with a loop.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON THE SUMMABILITY OF FOURIER SERIES†

BY
W. C. RANDELS

1. Notation. We shall use the following notation which is similar to that of Hille and Tamarkin‡ [2]. We shall use $f(x)$ to denote a function periodic of period 2π and integrable over $(-\pi, \pi)$. The Fourier series of $f(x)$ is

$$\sum_{n=-\infty}^{\infty} f_n e^{inx},$$

where

$$f_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

The class of such functions and their related Fourier series is denoted by L. We define

$$\tilde{f}(x) = \lim_{\epsilon\to 0}\left[-\frac{1}{2\pi}\int_{\epsilon}^{\pi}\{f(x+t)-f(x-t)\}\cot\frac{t}{2}\,dt\right] \tag{1.1}$$

and associate with it the series

$$-i\sum_{n=-\infty}^{\infty} \operatorname{sgn} n f_n e^{inx}.$$

We call the class of such series and their related functions $\tilde{L}$. If $f(x) \subset B.V.$ on $(-\pi, \pi)$, we denote by L' the class of functions $f'(x)$ and their related series,§

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d_t f(x+t) + i\sum_{n=-\infty}^{\infty} n f_n e^{inx}$$

and by $\tilde{L}'$ the class of functions

$$\tilde{f}'(x) = \lim_{\epsilon\to 0}\left[-\frac{1}{2\pi}\int_{\epsilon}^{\pi} d_t\{f(x+t)+f(x-t)-2f(x)\}\cot\frac{t}{2}\right] \tag{1.2}$$

and their related series

---

† Presented to the Society, January 2, 1936; received by the editors March 7, 1936.

‡ The numbers in brackets refer to the references at the end of the paper.

§ The Stieltjes integral used in this paper is the Young-Stieltjes integral over the open interval so that $\int_{-\pi}^{\pi} d_t f(x+t)$ need not be zero. This usage differs from that of Hille and Tamarkin [2].

$$\sum_{n=-\infty}^{\infty} | n | f_n e^{inx}.$$

We shall use the symbol $\mathfrak{T}$ for any one of these four classes. A function of class $\mathfrak{T}$ will be denoted by $F(x)$ and the general term of the associated series by $A_n(x)$.

We define

$$\phi(L, t) = f(x + t) + f(x - t) - 2f(x), \tag{1.3}$$

$$\phi(\tilde{L}, t) = f(x + t) - f(x - t), \tag{1.4}$$

$$\phi_0(L, t) = \int_0^t \phi(L, \tau)d\tau, \tag{1.5}$$

$$\phi_0(\tilde{L}, t) = \int_0^t \phi(\tilde{L}, \tau)d\tau, \tag{1.6}$$

$$\phi_0(L', t) = f(x + t) - f(x - t) - 2tf'(x), \tag{1.7}$$

$$\phi_0(\tilde{L}', t) = \phi(L, t). \tag{1.8}$$

We let $E(F, f)$ $(E(\tilde{F}, f))$ be the set of points x where $f(x)$ $(\tilde{f}(x))$ has a definite value and $\phi(L, t)\to 0$ as $t\to 0$ $(\phi(\tilde{L}, t)\to 0$ as $t\to 0)$. If $\phi_0(\mathfrak{T}, t) \subset B.V.$ on $(-\pi, \pi)$, we define $E(\mathfrak{T}, f)$ as the set of points where $F(x)$ has a definite value and†

$$\int_0^t | d\phi_0(\mathfrak{T}, \tau) | = o(t) \text{ as } t \to 0.$$

We now consider the transformation of a series defined by means of a matrix a_{mn} $(m, n=0, 1, \cdots)$ in the following manner

$$T_m(\mathfrak{T}, x) \sim \sum_{n=-\infty}^{\infty} a_{m|n|}A_n(x).$$

We make the restriction on a_{mn} that it define a regular method of summability, that is:

$$\sum_{n=0}^{\infty} | a_{mn} - a_{m,n+1} | < A, \quad A \text{ not depending on } m, \tag{1.9}$$

$$a_{mn} \to 1 \text{ as } m \to \infty, \text{ for every } n; \tag{1.10}$$

and that the series

$$\sum_{n=-\infty}^{\infty} a_{m|n|}A_n(x) \tag{1.11}$$

† By $\int_a^b | d\phi(t) |$ we mean the total variation of $\phi(t)$ on the interval (a, b).

be a Fourier series (cf. Hille and Tamarkin [3]). Then $T_m(\mathfrak{T}, x)$ is considered as a function of class L associated with (1.11). Thus $T_m(\mathfrak{T}, x)$ may be defined even though (1.11) does not converge. We denote the method of summability determined by the matrix a_{mn} by $\mathfrak{A}$.

If E_F stands for a set of points at which $F(x)$ has a definite value, we can make the following definitions:

DEFINITION 1. *A method of summation $\mathfrak{A}$ is said to be $(\mathfrak{T}, E_F)$-effective if, whenever $F(x) \subset \mathfrak{T}$ and $x \subset E_F$,*

$$T_m(\mathfrak{T}, x) - F(x) \to 0 \quad as \quad m \to \infty. \tag{1.12}$$

DEFINITION 2. *A method $\mathfrak{A}$ which is $(\mathfrak{T}, E_F)$-effective is said to be:*
(i) *F-effective, if $\mathfrak{T}=L$, $E_F = E(F, f)$,*
(ii) *$\tilde{F}$-effective, if $\mathfrak{T}=\tilde{L}$, $E_F = E(\tilde{F}, f)$,*
(iii) *L-effective, if $\mathfrak{T}=L$, $E_F = E(L, f)$,*
(iv) *$\tilde{L}$-effective, if $\mathfrak{T}=\tilde{L}$, $E_F = E(\tilde{L}, f)$,*
(v) *L'-effective, if $\mathfrak{T}=L'$, $E_F=E(L', f)$,*
(vi) *$\tilde{L}'$-effective, if $\mathfrak{T}=\tilde{L}'$, $E_F=E(\tilde{L}', f)$.*

2. $\mathfrak{T}=L$. By a theorem of Fekete [1] the most general factor sequence $\{a_n\}$ carrying a Fourier series into a Fourier series occurs when

$$\sum_{n=1}^{\infty} \frac{a_n}{n} \sin nx \subset B.V. \text{ on } (-\pi, \pi),$$

and

$$\sum_{n=1}^{\infty} a_n \{f_n e^{inx} + f_{-n} e^{-inx}\} \sim \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) d\left\{ \sum_{n=1}^{\infty} \frac{a_n}{n} \sin nt \right\}.$$

Then, if we let $\lim_{n\to\infty} a_{mn} = a_m$,

$$T_m(L, x) = a_{m0} f_0 + \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+t) d\left\{ \sum_{n=1}^{\infty} \frac{a_m}{n} \sin nt + \sum_{n=1}^{\infty} \frac{a_{mn} - a_m}{n} \sin nt \right\}.$$

But

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \begin{cases} \dfrac{\pi - x}{2}, & 0 < x < \pi, \\[2ex] \dfrac{-x-\pi}{2}, & -\pi < x < 0, \\[2ex] 0, & x = 0, \end{cases}$$

so that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)d\left\{\sum_{n=1}^{\infty}\frac{a_m}{n}\sin nt\right\} = a_m f(x) - a_m f_0.$$

Therefore for our problem we need only consider

$$(a_{m0} - a_m)f_0 + \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)d\left\{\sum_{n=1}^{\infty}\frac{a_{mn} - a_m}{n}\sin nt\right\}.$$

Hence it involves no loss of generality to suppose that $a_{mn}\to 0$ as $n\to\infty$, and we shall do so from now on.

Using Abel's method of partial summation we get

$$\begin{aligned}\sum_{n=1}^{\infty} a_{mn}\cos nx &= \sum_{n=1}^{\infty} a_{mn}\{S_n(x) - S_{n-1}(x)\}\\ &= -a_{m1}S_0(x) + \sum_{n=1}^{\infty}\{a_{mn} - a_{m,n+1}\}S_n(x),\end{aligned}$$

where

$$S_n(x) = \tfrac{1}{2} + \sum_{\nu=1}^{n}\cos\nu x = \frac{\sin(n+\frac{1}{2})x}{2\sin\frac{1}{2}x}.$$

Hence, by (1.9),

$$\sum_{n=1}^{\infty}\frac{a_{mn}}{n}\sin nx$$

is absolutely continuous on every interval not containing the origin. Moreover, again by Abel's transformation,

$$\sum_{n=1}^{\infty}\frac{a_{mn}}{n}\sin nx = \sum_{n=1}^{\infty}(a_{mn} - a_{m,n+1})\left(\sum_{\nu=1}^{n}\frac{1}{\nu}\sin\nu x\right)$$

and, since the partial sums

$$\sum_{\nu=1}^{n}\frac{1}{\nu}\sin\nu x$$

are bounded, the function

$$\sum_{n=1}^{\infty}\frac{a_{mn}}{n}\sin nx$$

is continuous at the origin and so, since it is of bounded variation by hypothesis, it must be absolutely continuous on the interval $(-\pi, \pi)$. Hence we can write

$$T_m(L, x) = a_{m0}f_0 + \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\left\{-\frac{a_{m1}}{2} + \sum_{n=1}^{\infty}(a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right\}dt$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\left\{\frac{a_{m0}}{2} - \frac{a_{m1}}{2} + \sum_{n=1}^{\infty}(a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right\}dt$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\left\{\sum_{n=0}^{\infty}(a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right\}dt,$$

and

$$T_m(L, x) - a_{m0}f(x) = \frac{1}{\pi}\int_0^{\pi}\phi(L, t)\left\{\sum_{n=0}^{\infty}(a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right\}dt.$$

By the Riemann-Lebesgue theorem

$$\int_0^{\pi}\phi(L, t)\left\{\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} - \frac{\sin nt}{t}\right\}dt = o(1) \text{ as } n \to \infty$$

so that, by the regularity of $\mathfrak{A}$,

$$T_m(L, x) - f(x) = \frac{1}{\pi}\int_0^{\pi}\phi(L, t)\left\{\sum_{n=0}^{\infty} a_{mn} - a_{m,n+1}\frac{\sin nt}{t}\right\}dt + o(1) \text{ as } m \to \infty.$$

For convenience we define

$$T_m = \frac{1}{\pi}\int_0^{\pi}\phi(L, t)\frac{1}{t}\sum_{n=0}^{\infty}(a_{mn} - a_{m,n+1})\sin nt\, dt.$$

We also define

$$q_m(u) = \sum_{0 \leq \nu < mu}(a_{m\nu} - a_{m,\nu+1}).$$

Then (1.9) can be written

$$\int_0^{\infty}|dq_m(u)| < A. \tag{2.1}$$

Then

$$T_m = \frac{1}{\pi}\int_0^{\pi}\phi(L, t)\frac{1}{t}H_m(mt)dt,$$

where

$$H_m(t) = \int_0^{\infty}\sin ut\, dq_m(u).$$

In this notation the familiar necessary and sufficient condition for F-effectiveness is that there exist an M such that

$$\int_0^\pi \left| \frac{1}{t} H_m(mt) \right| dt < M.$$

We wish to find a condition which will play a similar role for L-effectiveness. We introduce

$$H_m = \sum_{n=0}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} | H_m(t) |.$$

We shall consider only methods of summation for which there exists a constant B such that

$$| H_m(t) | < Bt. \tag{2.2}$$

This condition is always satisfied for methods defined by trapezoidal matrices since for such methods there is an N such that $q_m(u) = a_{m0}$ for $u > N$, and hence

$$| H_m(t) | = \left| \int_0^\infty \sin ut \, dq_m(u) \right| = \left| t \int_0^\infty [- a_{m0} + q(u)] \cos ut \, du \right| < ANt.$$

We now propose to show that if (2.2) is satisfied

$$T_m = o(H_m) + o(1) \tag{2.3}$$

and furthermore that this result is the best possible in the sense that, if $\overline{H}_m = \max_{\mu \leqq m} H_\mu \to \infty$ as $m \to \infty$, then for every sequence $\{d_m\}$ such that $d_m \downarrow 0$ as $m \to \infty$ it is possible to find a function $f(x)$ and a point $x \subset E(L, f)$ such that

$$T_m \neq O(d_m \overline{H}_m). \tag{2.4}$$

This implies that a necessary and sufficient condition for L-effectiveness of methods of summability satisfying (2.2) is the existence of an M such that

$$H_m < M. \tag{2.5}$$

In order to prove (2.3) we notice that by the Riemann-Lebesgue theorem for every a

$$\int_a^\pi \phi(L, t) \frac{\sin nt}{t} dt = o(1) \text{ as } n \to \infty,$$

so that by the regularity of $\mathfrak{A}$

$$\int_a^\pi \phi(L, t) \frac{1}{t} H_m(mt) dt = o(1) \text{ as } m \to \infty. \tag{2.6}$$

Moreover, if $x \subset E(L, f)$,

$$
\begin{aligned}
\left|\int_0^a \phi(L, t) \frac{1}{t} H_m(mt)dt\right|
& \leqq \left|\int_0^{1/m} \phi(L, t) \frac{1}{t} H_m(mt)dt\right| + \left|\int_{1/m}^a \phi(L, t) \frac{1}{t} H_m(mt)dt\right| \\
& < Bm \int_0^{1/m} |\phi(L, t)| dt + \left|\int_1^{ma} \phi(L, t/m) \frac{1}{t} H_m(t)dt\right| \qquad (2.7) \\
& = o(1) + O\left(\sum_{n=0}^{[\log_2 ma]} \max_{2^n \leqq t \leqq 2^{n+1}} |H_m(t)| \, 2^{-n} \int_1^{2^{n+1}} |\phi(L, t/m)| \, dt\right) \\
& = o(1) + o(H_m) \text{ uniformly in } m \text{ as } a \to 0.
\end{aligned}
$$

Then (2.3) follows from (2.6) and (2.7) by a familiar technique.

In order to show that this result is the best possible it is sufficient to prove that for every fixed a it is possible to find a function $f(x)$ and an integer m such that

$$\frac{1}{\pi} \int_{1/m}^a f(t) \frac{1}{t} H_m(mt)dt > \overline{H}_m \tag{2.8}$$

and there is a constant K independent of a such that

$$\int_{1/m}^t |f(t)| \, dt < Kt, \quad t \leqq a. \tag{2.9}$$

For, if $d_m \downarrow 0$ is given, then making use of the Riemann-Lebesgue theorem and the regularity of $\mathfrak{A}$ we construct by induction a sequence of integers $\{m_i\}$ and functions $f_i(x)$ such that $m_0 = 1$, and

$$\sum_{i=1}^{\infty} (d_{m_i})^{1/2} < \infty, \tag{2.10}$$

$$\left| \frac{1}{\pi} \int_{1/m_j}^{1/m_{j-1}} f_i(t) \frac{1}{t} H_{m_i}(m_i t)dt \right| < 1, \qquad j < i, \tag{2.11}$$

$$\frac{1}{\pi} \int_{1/m_i}^{1/m_{i-1}} f_i(t) \frac{1}{t} H_{m_i}(m_i t)dt > \overline{H}_{m_i}, \tag{2.12}$$

and

$$\frac{1}{\pi} \int_{1/m_i}^t |f_i(t)| \, dt < Kt, \qquad t \leqq \frac{1}{m_{i-1}}. \tag{2.13}$$

We define

$$f(x) = \begin{cases} (d_{m_i})^{1/2} f_i(|x|), & \dfrac{1}{m_i} \leqq |x| < \dfrac{1}{m_{i-1}}, \\ 0, \quad x = 0, & 1 \leqq |x| \leqq \pi. \end{cases}$$

Then at $x=0$, $\phi(L, t) = 2f(t)$ and, by (2.13), if $1/m_i \leqq t \leqq 1/m_{i-1}$,

$$\int_0^t |\phi(L, t)| \, dt = 2 \sum_{j=i+1}^{\infty} (d_{m_j})^{1/2} \int_{1/m_j}^{1/m_{j-1}} |f_j(t)| \, dt + 2(d_{m_i})^{1/2} \int_{1/m_i}^{t} |f_i(t)| \, dt$$

$$< 2Kt \sum_{j=i}^{\infty} (d_{m_j})^{1/2} = o(t) \text{ as } t \to 0,$$

so that the point $x=0$ is in $E(L, f)$. Moreover

$$T_{m_i} = \frac{1}{\pi} \int_0^{\pi} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt$$

$$= \frac{1}{\pi} \int_0^{1/m_i} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt + \frac{1}{\pi} \int_{1/m_i}^{1/m_{i-1}} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt$$

$$+ \frac{1}{\pi} \int_{1/m_{i-1}}^{\pi} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt.$$

By (2.2)

$$\left| \frac{1}{\pi} \int_0^{1/m_i} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt \right| < \frac{Bm_i}{\pi} \int_0^{1/m_i} |\phi(L, t)| \, dt = o(1)$$

and by (2.11)

$$\frac{1}{\pi} \int_{1/m_{i-1}}^{\pi} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt = \sum_{j=1}^{i-1} \frac{2(d_{m_j})^{1/2}}{\pi} \int_{1/m_j}^{1/m_{j-1}} f_j(t) \frac{1}{t} H_{m_i}(m_i t) dt$$

$$= O\left(\sum_{j=1}^{\infty} (d_{m_j})^{1/2} \right) = O(1).$$

Finally by (2.12)

$$\frac{1}{\pi} \int_{1/m_i}^{1/m_{i-1}} \phi(L, t) \frac{1}{t} H_{m_i}(m_i t) dt$$

$$= \frac{2(d_{m_i})^{1/2}}{\pi} \int_{1/m_i}^{1/m_{i-1}} f(t) \frac{1}{t} H_{m_i}(m_i t) dt > 2(d_{m_i})^{1/2} \overline{H}_{m_i},$$

so that

$$T_{m_i} > 2(d_{m_i})^{1/2} \overline{H}_{m_i} - o(1) - O(1) \neq O(d_{m_i} \overline{H}_{m_i}).$$

It only remains now to prove the existence of an $f(x)$ and an m with the properties (2.8), (2.9). We notice that

$$|H_m(t)| = \left|\int_0^\infty \sin ut\, dq_m(u)\right| < \int_0^\infty |dq_m(u)| < A,$$

so that for a fixed $a<1$

$$\sum_{n=0}^{[\log_2 ma]-1} \max_{2^n \leqq t \leqq 2^{n+1}} |H_m(t)| > \sum_{n=0}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} |H_m(t)| - A(|\log a| + 1).$$

Then since $\overline{H}_m \uparrow \infty$ as $m \to \infty$, we can find an m so that $\overline{H}_m = H_m$ and m also may be chosen so that

$$\sum_{n=0}^{[\log_2 ma]-1} \max_{2^n \leqq t \leqq 2^{n+1}} |H_m(t)| > \tfrac{1}{2}\overline{H}_m.$$

We call the point on $(2^n, 2^{n+1})$ where $H_m(t)$ has maximum absolute value x_{mn}. The function $H_m(t)$ is continuous and hence there will be an interval J_{mn} containing x_{mn} and contained in $(2^n, 2^{n+1})$ such that for $x \subset J_{mn}$

$$|H_m(x_{mn}) - H_m(x)| < \tfrac{1}{2}|H_m(x_{mn})|. \tag{2.14}$$

Then J^*_{mn} is defined as the interval into which J_{mn} is carried by the transformation $x' = x/m$. It is clear that J^*_{mn} is contained in $(2^n/m, 2^{n+1}/m)$. We denote the length of J^*_{mn} by ϵ_{mn}. We define

$$f(x) = \begin{cases} \dfrac{4\pi x \operatorname{sgn} H_m(x_{mn})}{\epsilon_{mn}}, & x \subset J^*_{mn}, \quad 0 \leqq n \leqq \log_2 ma - 1, \\[2ex] 0, & \text{elsewhere on } \left(\dfrac{1}{m}, a\right). \end{cases}$$

The function $f(x)$ satisfies (2.9), for, if $2^n/m \leqq t < 2^{n+1}/m$ $(0 \leqq n \leqq \log_2 ma - 1)$,

$$\int_{1/m}^t |f(x)|\, dx \leqq \int_{1/m}^{2^{n+1}/m} |f(x)|\, dx \leqq 4\pi \sum_{\nu=1}^{n+1} \frac{2^\nu}{m} < \frac{4\pi(2^{n+1}-1)}{m} < 8\pi t.$$

By (2.14)

$$\begin{aligned} \frac{1}{\pi}\int_{1/m}^a f(x)\frac{1}{x} H_m(mx)dx &= \frac{1}{\pi}\sum_{\nu=0}^{[\log_2 ma]-1} \int_{J^*_{mn}} f(x)\frac{1}{x}H_m(mx)dx \\ &= \frac{1}{\pi}\sum_{\nu=0}^{[\log_2 ma]-1} \frac{4\pi \operatorname{sgn} H(x_{mn})}{\epsilon_{mn}} \int_{J^*_{mn}} H_m(mx)dx \\ &> \sum_{n=0}^{[\log_2 ma]-1} \frac{4}{\epsilon_{mn}}\epsilon_{mn}\left|\frac{H_m(x_{mn})}{2}\right| > 2\sum_{n=0}^{[\log_2 ma]-1} |H_m(x_{mn})| > \overline{H}_m, \end{aligned}$$

so that (2.8) is satisfied and our proof is completed.

3. $\mathfrak{T}=\tilde{L}$. Fekete [1] has shown that a necessary and sufficient condition that a factor sequence a_n carry a series of $\tilde{L}$ into a Fourier series is that

$$\sum_{n=1}^{\infty} \frac{a_n}{n} \cos nt \subset B.V. \text{ on } (-\pi, \pi);$$

and

$$-i\sum_{n=1}^{\infty} a_n\{f_n e^{inx} - f_{-n}e^{-inx}\} \sim \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)d\left\{\sum_{n=1}^{\infty} \frac{a_n}{n} \cos nt\right\}.$$

Therefore

$$T_m(L, x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)d\left\{\sum_{n=1}^{\infty} \frac{a_{mn}}{n} \cos nt\right\}.$$

We make the hypothesis that $a_{mn}\to 0$ as $n\to\infty$. This condition is essential here because

$$\sum_{n=1}^{\infty} \frac{1}{n} \cos nt \not\subset B.V. \text{ on } (-\pi, \pi)$$

and therefore the argument of §2 does not apply. Then

$$\sum_{n=1}^{\infty} a_{mn} \sin nt = \sum_{n=1}^{\infty} a_{mn}\{\tilde{S}_n(t) - \tilde{S}_{n-1}(t)\} = \sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1})\tilde{S}_n(t),$$

where

$$\tilde{S}_n(t) = \sum_{\nu=1}^{n} \sin \nu t = -\frac{1}{2}\cot\frac{t}{2} + \frac{\cos (n+\frac{1}{2})t}{2\sin\frac{1}{2}t}.$$

Hence, by (1.9),

$$g(t) = \sum_{n=1}^{\infty} \frac{a_{mn}}{n} \cos nt$$

is absolutely continuous on every interval not containing the origin and, since $g(-0)=g(+0)$, $g(t)$ can be considered as continuous at the origin so that, since by hypothesis $g(t)$ is of bounded variation, it must be absolutely continuous on $(-\pi, \pi)$. Then

$$\begin{aligned} T_m(\tilde{L}, x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1})\frac{-\cos\frac{1}{2}t + \cos (n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\, dt \\ &= \frac{1}{\pi}\int_{0}^{\pi} \phi(\tilde{L}, t)\sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1})\frac{-\cos\frac{1}{2}t + \cos (n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\, dt. \end{aligned}$$

By the Riemann-Lebesgue theorem

$$\int_0^\pi \phi(\tilde{L}, t)\left\{\frac{\cos (n+\frac{1}{2})t}{2 \sin \frac{1}{2}t} - \frac{\cos nt}{t}\right\} dt = o(1) \text{ as } n \to \infty,$$

so that by the regularity of $\mathfrak{A}$

$$T_m(\tilde{L}, x) = \frac{1}{\pi}\int_0^\pi \phi(\tilde{L}, t)\left\{\frac{-a_{m1}}{2} \cot \frac{t}{2} + \frac{1}{t}\tilde{H}_m(mt)\right\} dt + o(1) \text{ as } m \to \infty$$

where

$$\tilde{H}_m(t) = \int_0^\infty \cos ut\, dq_m(u).$$

By the definition of $f(x)$ and the regularity of $\mathfrak{A}$

$$-\frac{a_{m1}}{2\pi}\int_{1/m}^\pi \phi(\tilde{L}, t) \cot \frac{t}{2}\, dt \to \tilde{f}(x) \text{ as } m \to \infty,$$

if $x \subset E(\tilde{L}, f)$, and also,

$$\int_0^{1/m} \phi(\tilde{L}, t)\left\{\frac{1}{2} \cot \frac{t}{2} - \frac{1}{t}\right\} dt = o(1) \text{ as } m \to \infty,$$

so that

$$T_m(\tilde{L}, x) - a_{m1}\tilde{f}(x) = \frac{1}{\pi}\int_0^{1/m} \phi(\tilde{L}, t)\,\frac{1}{t}\,\{-a_{m1} + \tilde{H}_m(mt)\}\, dt$$
$$+ \frac{1}{\pi}\int_{1/m}^\pi \phi(\tilde{L}, t)\,\frac{1}{t}\,\tilde{H}_m(mt)\,dt + o(1) \text{ as } m \to \infty.$$

We make a hypothesis analogous to (2.2), namely,

$$|a_{m1} - \tilde{H}_m(t)| < Bt. \tag{3.1}$$

Then, if $x \subset E(\tilde{L}, f)$,

$$\left|\frac{1}{\pi}\int_0^{1/m} \phi(\tilde{L}, t)\,\frac{1}{t}\,\{-a_{m1} + \tilde{H}_m(mt)\}\, dt\right|$$
$$< \frac{Bm}{\pi}\int_0^{1/m} |\phi(\tilde{L}, t)|\, dt = o(1) \text{ as } m \to \infty.$$

We shall prove that, if $x \subset E(\tilde{L}, f)$,

$$\tilde{T}_m = \frac{1}{\pi}\int_{1/m}^\pi \phi(\tilde{L}, t)\,\frac{1}{t}\,\tilde{H}_m(mt)\,dt = o(\tilde{H}_m) + o(1), \tag{3.2}$$

where

$$\tilde{\tilde{H}}_m = \sum_{n=0}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} | \tilde{H}_m(t) |.$$

If a is fixed, by the Riemann-Lebesgue theorem and the regularity of $\mathfrak{A}$,

$$\frac{1}{\pi}\int_a^\pi \phi(\tilde{L}, t)\frac{1}{t}\tilde{H}_m(mt)dt = o(1) \text{ as } m \to \infty,$$

and

$$\begin{aligned}\frac{1}{\pi}\int_{1/m}^a \phi(\tilde{L}, t)\frac{1}{t}\tilde{H}_m(mt)dt &= \frac{1}{\pi}\int_1^{ma} \phi\left(\tilde{L}, \frac{t}{m}\right)\frac{1}{t}\tilde{H}_m(t)dt \\ &= O\left(\sum_{n=0}^{[\log_2 m]} \max_{2^n\leqq t\leqq 2^{n+1}} | \tilde{H}_m(t) | \, 2^{-n}\int_0^{2^{n+1}} \left|\phi\left(\tilde{L}, \frac{t}{m}\right)\right| dt\right) \\ &= o(\tilde{\tilde{H}}_m) \text{ uniformly in } m \text{ as } a \to 0,\end{aligned}$$

so that (3.2) must be satisfied.

We would like to prove that this result is the best possible. The method of §2 however does not seem to be capable of doing this without some additional assumption. We choose to assume that

$$\int_{1/m}^\pi \frac{1}{t} | \tilde{H}_m(mt) | \, dt < K. \tag{3.3}$$

This condition is not completely arbitrary for it is analogous to the necessary and sufficient condition for F-effectiveness mentioned in §2. It is known to be sufficient for $\tilde{F}$-effectiveness but has not been proved to be necessary. It will be seen from the proof that some slightly less restrictive condition such as

$$\sum_{n=0}^{[\log_2 m]} \min_{2^n\leqq t\leqq 2^{n+1}} | H_m(t) | < k\tilde{\tilde{H}}_m, \qquad k < 1,$$

could be used.

We define $\tilde{\tilde{\tilde{H}}}_m = \max_{\mu\leqq m}\tilde{\tilde{H}}_\mu$ and we wish to show that, if $\tilde{\tilde{\tilde{H}}}_m \uparrow \infty$ as $m \to \infty$, and $\{d_m\}$ is any sequence such that $d_m \downarrow 0$ as $m \to \infty$, we can find a function $f(x)$ and a point $x \subset E(\tilde{L}, f)$ for which

$$\tilde{T}_m \neq O(d_m \tilde{\tilde{\tilde{H}}}_m).$$

As in §2 it is sufficient to show that for a fixed it is possible to find a function $f(x)$ and an integer m for which

$$(3.4)\qquad \frac{1}{\pi}\int_{1/m}^{a} f(x)\,\frac{1}{x}\,\tilde{H}_m(mx)dx > \tilde{\tilde{H}}_m,$$

$$(3.5)\qquad \int_{1/m}^{t} |\,f(x)\,|\,dx < Ct, \qquad C \text{ not depending on } a \text{ or } m,$$

and

$$(3.6)\qquad \left|\int_{\epsilon}^{a} \frac{f(x)}{x}\,dx\right| < D, \qquad D \text{ not depending on } a \text{ or } \epsilon.$$

For, following the procedure of §2, we construct a sequence of functions $f_i(x)$ and integers m_i such that $m_0=1$ and

$$(3.7)\qquad \sum_{i=1}^{\infty} (d_{m_i})^{1/2} < \infty,$$

$$(3.8)\qquad \left|\frac{1}{\pi}\int_{1/m_j}^{1/m_{j-1}} f_i(x)\,\frac{1}{x}\,\tilde{H}_{m_i}(m_i x)dx\right| < 1, \qquad j < i,$$

$$(3.9)\qquad \frac{1}{\pi}\int_{1/m_i}^{1/m_{i-1}} f_i(x)\,\frac{1}{x}\,\tilde{H}_{m_i}(m_i x)dx > \tilde{\tilde{H}}_{m_i},$$

$$(3.10)\qquad \int_{1/m_i}^{t} |\,f_i(x)\,|\,dx < Ct, \qquad \frac{1}{m_i} \leqq t \leqq \frac{1}{m_{i-1}},$$

and

$$(3.11)\qquad \left|\int_{\epsilon}^{1/m_{i-1}} \frac{f_i(x)}{x}\,dx\right| < D, \qquad \frac{1}{m_i} \leqq \epsilon \leqq \frac{1}{m_{i-1}},$$

Then $f(x)$ is defined by

$$f(x) = \begin{cases} (d_{m_i})^{1/2}\,\mathrm{sgn}\,x f_i(|\,x\,|), & -\dfrac{1}{m_{i-1}} < x \leqq -\dfrac{1}{m_i},\quad \dfrac{1}{m_i} \leqq x < \dfrac{1}{m_{i-1}} \\ & \qquad\qquad (i = 1, 2, \cdots) \\ 0, & -\pi \leqq x \leqq -1, \quad x = 0, \quad 1 \leqq x \leqq \pi. \end{cases}$$

At $x=0$, $\phi(\tilde{L}, t)=2f(t)$ and, if $1/m_i \leqq t < 1/m_{i-1}$,

$$\int_0^t |\,\phi(\tilde{L}, t)\,|\,dt = 2\sum_{j=i+1}^{\infty} (d_{m_j})^{1/2}\int_{1/m_j}^{1/m_{j-1}} |\,f_j(x)\,|\,dx + 2(d_{m_i})^{1/2}\int_{1/m_i}^{t} |\,f_i(x)\,|\,dx$$

$$< 2Ct\sum_{j=i}^{\infty} (d_{m_j})^{1/2} = o(t) \text{ as } t \to 0.$$

Since $(\cot t/2 - 2/t)$ is a bounded function, the existence of

$$\tilde{f}(x) = \lim_{\epsilon\to 0}\left[-\frac{1}{2\pi}\int_\epsilon^\pi \phi(\tilde{L}, t)\cot\frac{t}{2}\,dt\right],$$

is equivalent to the existence of

$$\lim_{\epsilon\to 0}\left[-\frac{1}{\pi}\int_\epsilon^\pi \phi(\tilde{L}, t)\frac{dt}{t}\right].$$

By (3.11)

$$\lim_{\epsilon\to 0}\left[-\frac{1}{\pi}\int_\epsilon^\pi \phi(\tilde{L}, t)\frac{dt}{t}\right] = -\sum_{j=1}^{\infty}\frac{2}{\pi}(d_{m_j})^{1/2}\int_{1/m_j}^{1/m_{j-1}} f_j(x)\frac{dx}{x}$$

and by (3.7) and (3.11) the second sum converges. Therefore the point $x=0$ is contained in $E(\tilde{L}, f)$. Finally

$$\tilde{T}_{m_i} = \int_{1/m_i}^{1/m_{i-1}} \phi(\tilde{L}, t)\tilde{H}_{m_i}(m_i t)dt + \int_{1/m_{i-1}}^{\pi} \phi(\tilde{L}, t)\tilde{H}_{m_i}(m_i t)dt$$

$$> 2(d_{m_i})^{1/2}\tilde{\tilde{H}}_{m_i} - O\left(\sum_{j=1}^{i-1}(d_{m_j})^{1/2}\right) \neq O(d_{m_i}\tilde{\tilde{H}}_{m_i}).$$

We now need only construct a function $f(x)$ and an integer m with the properties (3.4), (3.5), and (3.6). First

$$|\tilde{H}_m(t)| = \left|\int_0^\infty \cos ut\, dq_m(u)\right| < A,$$

so that

$$\sum_{n=0}^{[\log_2 ma]-1}\max_{2^n\leq t\leq 2^{n+1}}|\tilde{H}_m(t)| > \sum_{n=0}^{[\log_2 m]}\max_{2^n\leq t\leq 2^{n+1}}|\tilde{H}_m(t)| - A(|\log a| + 1),$$

and therefore if $\tilde{\tilde{H}}_m \uparrow \infty$ as $m\to\infty$, we can find an m so that

$$\sum_{n=0}^{[\log_2 ma]-1}\max_{2^n\leq t\leq 2^{n+1}}|\tilde{H}_m(t)| - 2K > \tfrac{1}{2}\tilde{\tilde{H}}_m,$$

where K is the constant of condition (3.3). We call the point on $(2^n, 2^{n+1})$ where $\tilde{H}_m(t)$ attains its maximum absolute value x_{mn}. Since $\tilde{H}_m(t)$ is continuous there must be an interval J_{mn} containing x_{mn} and contained in $(2^n, 2^{n+1})$ such that for $x \subset J_{mn}$

$$|\tilde{H}_m(x_{mn}) - \tilde{H}_m(x)| < \tfrac{1}{2}\tilde{H}_m(x_{mn}).$$

We call the point on $(2^n, 2^{n+1})$ where $\tilde{H}_m(t)$ attains its minimum absolute value $\bar{x}_{mn}$. Then (3.3) implies that

$$\sum_{n=0}^{[\log_2 m]} |\tilde{H}_m(\bar{x}_{mn})| < K.$$

We can find an interval G_{mn} containing $\bar{x}_{mn}$ and contained in $(2^n, 2^{n+1})$ such that for $x \subset G_{mn}$

$$|\tilde{H}_m(\bar{x}_{mn}) - \tilde{H}_m(x)| < \tfrac{1}{8}\tilde{H}_m(x_{mn}).$$

We may suppose without loss of generality that the lengths of J_{mn} and G_{mn} are equal. We denote by J^*_{mn} (G^*_{mn}) the image of J_{mn} (G_{mn}) under the transformation $x' = x/m$ and we call the length of J^*_{mn}, ϵ_{mn}. Then $f(x)$ is defined by

$$f(x) = \begin{cases} \dfrac{8\pi x \operatorname{sgn} \tilde{H}_m(x_{mn})}{\epsilon_{mn}}, & x \subset J^*_{mn} \quad (n = 0, 1, \cdots, [\log_2 ma] - 1), \\ -\dfrac{8\pi x \operatorname{sgn} \tilde{H}_m(x_{mn})}{\epsilon_{mn}}, & x \subset G^*_{mn} \quad (n = 0, 1, \cdots, [\log_2 ma] - 1), \\ 0, \quad \text{elsewhere on } \left(\dfrac{1}{m}, a\right). \end{cases}$$

The function $f(x)$ satisfies (3.4), for

$$\begin{aligned} &\frac{1}{\pi}\int_{1/m}^{a} f(x)\frac{1}{x}\tilde{H}_m(mx)dx \\ &= \frac{1}{\pi}\sum_{n=0}^{[\log_2 ma]-1}\left\{\int_{J^*_{mn}} f(x)\frac{1}{x}\tilde{H}_m(mx)dx + \int_{G^*_{mn}} f(x)\frac{1}{x}\tilde{H}_m(mx)dx\right\} \\ &> 4\sum_{n=0}^{[\log_2 ma]-1} |\tilde{H}_m(x_{mn})| - 8\sum_{n=0}^{[\log_2 ma]-1} \max_{x\subset G^*_{mn}} |\tilde{H}_m(mx)| \\ &> 4\sum_{n=0}^{[\log_2 ma]-1} |\tilde{H}_m(x_{mn})| - 8\sum_{n=0}^{[\log_2 ma]-1} |\tilde{H}_m(\bar{x}_{mn})| - 8\sum_{n=0}^{[\log_2 ma]-1} \tfrac{1}{8}|\tilde{H}_m(x_{mn})| \\ &> 2\tilde{\tilde{H}}_m - \tilde{\tilde{H}}_m = \tilde{\tilde{H}}_m. \end{aligned}$$

If $2^n/m \leqq t < 2^{n+1}/m$, $0 \leqq n \leqq [\log_2 ma] - 1$, we have

$$\int_{1/m}^{t} |f(x)|\,dx \leqq 8\pi\sum_{\nu=1}^{n+1}\frac{2^\nu}{m} < \frac{8\pi 2^{n+1}}{m} < 16\pi t,$$

and (3.5) is satisfied. If $2^n/m \leqq \epsilon < 2^{n+1}/m$, $(n = 0, 1, \cdots, [\log_2 ma] - 1)$,

$$\int_{\epsilon}^{a} f(x)\frac{dx}{x} = \sum_{\nu=n+1}^{[\log_2 ma]-1}\left\{\int_{J_{mn}^{*}} \frac{8\pi \operatorname{sgn} \tilde{H}_m(x_{mn})}{\epsilon_{mn}}dx - \int_{G_{mn}^{*}} \frac{8\pi \operatorname{sgn} \tilde{H}_m(x_{mn})}{\epsilon_{mn}}dx\right\}$$
$$+ \int_{2^n/m}^{2^{n+1}/m} f(x)\frac{dx}{x} \leqq 8\pi$$

which proves (3.6) and completes the proof of the theorem.

4. $\mathfrak{T}=L'$. In this case we again make the restriction that $a_{mn}\to 0$ as $n\to\infty$. This involves no loss of generality for, if $T_m(L', x) \subset L$, then we must have $na_{mn}f_n\to 0$ as $n\to\infty$, whenever $f'(x)$ is in L'. This can only be the case if $a_{mn}\to 0$ as $n\to\infty$. We have

$$T_m(L', f) = \frac{1}{\pi}\int_{-\pi}^{\pi} d_t f(x+t)\sum_{n=0}^{\infty}(a_{mn}-a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t},$$

where

$$\sum_{n=0}^{\infty}(a_{mn}-a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \subset L \text{ on } (-\pi, \pi).$$

Hence, if $x \subset E(L', f)$,

$$T_m(L', f) - a_{m1}f'(x) = \frac{1}{\pi}\int_{0}^{\pi} d\phi_0(L', t)\sum_{n=1}^{\infty}(a_{mn}-a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}.$$

We wish to determine a necessary and sufficient condition for L'-effectiveness when the conditions

$$(4.1) \qquad |H_m(t)| < Bt,$$

$$(4.2) \qquad \sum_{n=0}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} |H_m(t)| < M$$

are satisfied. The required condition turns out to be

$$(4.3) \qquad \sum_{n=0}^{\infty}(a_{mn}-a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \to 0 \text{ as } m\to\infty \text{ for } t>0.$$

This condition is necessary for, if it is not satisfied, let $t_0>0$ be a point where

$$\overline{\lim_{m\to\infty}}\left|\sum_{n=0}^{\infty}(a_{mn}-a_{m,n+1})\frac{\sin(n+\frac{1}{2})t_0}{2\sin\frac{1}{2}t_0}\right| > 0.$$

Then, if $f(x)$ is defined by

$$f(x) = \begin{cases}\operatorname{sgn} x, & t_0 \leqq |x| < \pi,\\ 0, & 0 \leqq |x| < t_0, \qquad x = \pi,\end{cases}$$

at $x=0$, $\phi(L', t)=2f(t)$ and $f'(0)=0$,

$$\int_0^t | d\phi_0(L', \tau) | = 0, \qquad t \leqq t_0,$$

so that the point $x=0$ is contained in $E(L',f)$. But

$$\begin{aligned} T_m(L', x) - a_{m1}f'(x) &= \frac{1}{\pi}\int_0^\pi d\phi_0(L',t)\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \\ &= \frac{1}{\pi}\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t_0}{2\sin\frac{1}{2}t_0} \end{aligned}$$

which does not tend to zero as $m\to\infty$ so the method cannot be L'-effective.

The condition (4.3) is sufficient, for

$$\begin{aligned} T_m(L', x) - a_{m1}f'(x) &= \frac{1}{\pi}\int_0^a d\phi_0(L',t)\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \\ &+ \frac{1}{\pi}\int_a^\pi d\phi_0(L',t)\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}. \end{aligned}$$

Since $\sin((n+\frac{1}{2})t)/(2\sin(\frac{1}{2}t))-(\sin nt)/t$ is bounded, if $x \subset E(L',f)$,

$$\begin{aligned} &\frac{1}{\pi}\int_0^a d\phi_0(L',t)\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \\ &\qquad = \frac{1}{\pi}\int_0^a d\phi_0(L',t)\frac{1}{t}H_m(mt) + o(1) \text{ uniformly in } m \text{ as } a\to 0, \end{aligned}$$

and, by (4.1) and (4.2),

$$\frac{1}{\pi}\int_0^{1/m} d\phi_0(L',t)\frac{1}{t}H_m(mt) < \frac{Bm}{\pi}\int_0^{1/m} | d\phi_0(L',t) | = o(1) \text{ as } m\to\infty,$$

$$\begin{aligned} &\frac{1}{\pi}\int_{1/m}^a d\phi_0(L',t)\frac{1}{t}H_m(mt) \\ &\qquad = O\left(\sum_{n=0}^{[\log_2 ma]} \max_{2^n\leqq t\leqq 2^{n+1}} | H_m(t) | 2^{-n}m\int_0^{2^{n+1}} | d\phi_0(L', t/m) |\right) \\ &\qquad = o(1) \text{ uniformly in } m \text{ as } a\to 0. \end{aligned}$$

Finally, if (4.3) is satisfied, since

$$\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} = O\left(\frac{A}{t}\right),$$

$$\frac{1}{\pi}\int_a^\pi d\phi_0(L',t)\sum_{n=0}^\infty (a_{mn} - a_{m,n+1})\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} = o(1) \text{ as } m\to\infty.$$

By putting these results together we have the proof of sufficiency.

5. $\mathfrak{T}=\tilde{L}'$. In this case as in the case considered in §4 the condition $a_{mn}\to 0$ as $n\to\infty$ is essential in order that $T_m(\tilde{L}', x) \subset L$. Then

$$\begin{aligned} T_m(\tilde{L}', x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} d_t f(x+t)\sum_{n=1}^{\infty} a_{mn} \sin nt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} d_t f(x+t)\sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1}) \frac{-\cos\frac{1}{2}t + \cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \\ &= \frac{1}{\pi}\int_{0}^{\pi} d\phi_0(\tilde{L}', t)\sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1}) \frac{-\cos\frac{1}{2}t + \cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}. \end{aligned}$$

We shall assume that

$$(5.1) \qquad | a_{m1} - \tilde{H}_m(t) | < Bt$$

and

$$(5.2) \qquad \sum_{n=1}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} | \tilde{H}_m(t) | < M,$$

and we propose to find a necessary and sufficient condition for $\tilde{L}'$-effectiveness. This condition is

$$(5.3) \qquad \sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1}) \frac{\cos(n+\frac{1}{2})t}{\sin\frac{1}{2}t} \to 0 \text{ as } m \to \infty, \text{ for } t > 0.$$

The condition is necessary for, if it is not satisfied, let $t_0>0$ be a point such that

$$\overline{\lim_{m\to\infty}} \left| \sum_{n=1}^{\infty} (a_{mn} - a_{m,n+1}) \frac{\cos(n+\frac{1}{2})t_0}{\sin\frac{1}{2}t_0} \right| > 0.$$

Then we define

$$f(x) = \begin{cases} 1, & -\pi \leqq x \leqq -t_0, \quad t_0 \leqq x \leqq \pi; \\ 0, & -t_0 < x < t_0. \end{cases}$$

For this function at $x=0$, $\phi_0(\tilde{L}', t)=2f(t)$ and

$$\tilde{f}'(0) = \frac{-1}{2\pi}\int_0^{\pi} d\phi_0(\tilde{L}', t) \cot\frac{t}{2} = -\frac{1}{\pi}\cot\frac{t_0}{2}.$$

Also

$$\int_0^t | d\phi_0(\tilde{L}', t) | = 0, \text{ for } t < t_0,$$

so that the point $x=0$ is contained in $E(\tilde{L}',f)$. But

$$T_m(\tilde{L}', x) = \frac{1}{\pi}\int_0^\pi d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{-\cos\frac{1}{2}t + \cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}$$
$$= a_{m1}\tilde{f}'(0) + \frac{1}{\pi}\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{\cos(n+\frac{1}{2})t_0}{\sin\frac{1}{2}t_0}$$

which cannot tend to $\tilde{f}'(0)$ as $m\to\infty$ and hence the method is not $\tilde{L}'$-effective.

If (5.3) is satisfied and $x \subset E(\tilde{L}',f)$,

$$\frac{1}{\pi}\int_a^\pi d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{\cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} \to 0 \text{ as } m\to\infty,$$

$$-\frac{1}{2\pi}\int_{1/m}^\pi d\phi_0(\tilde{L}', t)a_{m1}\cot\frac{t}{2} \to \tilde{f}'(x) \text{ as } m\to\infty,$$

$$\frac{1}{\pi}\int_0^{1/m} d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{-\cos\frac{1}{2}t + \cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}$$
$$= \frac{1}{\pi}\int_0^{1/m} d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{-1+\cos nt}{t} + O\left(\int_0^{1/m} | d\phi_0(\tilde{L}', t) |\right)$$
$$= O\left(\frac{m}{\pi}\int_0^{1/m} | d\phi_0(\tilde{L}', t) |\right) + o(1) = o(1) \text{ as } m\to\infty,$$

and

$$\frac{1}{\pi}\int_{1/m}^a d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{\cos(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}$$
$$= \frac{1}{\pi}\int_{1/m}^a d\phi_0(\tilde{L}', t)\sum_{n=1}^\infty (a_{mn} - a_{m,n+1})\frac{\cos nt}{t} + o(1)$$
$$= O\left(\sum_{n=1}^{[\log_2 ma]} \max_{2^n\le t\le 2^{n+1}} | \tilde{H}_m(t) | 2^{-n}m\int_0^{2^{n+1}} | d\phi_0(\tilde{L}', t/m) |\right) + o(1)$$
$$= o(1) \text{ uniformly in } m \text{ as } a\to\infty.$$

Therefore conditions (5.1), (5.2), and (5.3) are sufficient for $\tilde{L}'$-effectiveness.

6. **A necessary condition.** We shall consider methods of summability for which

$$\int_N^\infty | dq_m(u) | \to 0 \text{ uniformly in } m \text{ as } N\to\infty \tag{6.1}$$

and

(6.2) $\qquad q_m(u) \to q(u)$ as $m \to \infty$, over an everywhere dense set.

The definition of $q(u)$ could be extended to the whole interval $(0, \infty)$ by setting

$$q(u) = \tfrac{1}{2}\{q(u+0) + q(u-0)\}$$

and $q(u)$ would have the property

$$\int_0^\infty |\,dq(u)\,| < A.$$

Under these assumptions we have

$$H_m(t) \to H(t) = \int_0^\infty \sin ut\, dq(u)$$

for

$$|\,H_m(t) - H(t)\,| \leqq \left| \int_0^N \sin ut\, d\{q_m(u) - q(u)\} \right| + \int_N^\infty |\,dq_m(u)\,| + \int_N^\infty |\,dq(u)\,|$$

and, by (6.1),

$$\int_N^\infty |\,dq(u)\,| + \int_N^\infty |\,dq_m(u)\,| \to 0 \text{ uniformly in } m \text{ as } N \to \infty,$$

and for N fixed

$$\begin{aligned} &\left| \int_0^N \sin ut\, d\{q_m(u) - q(u)\} \right| \\ &\qquad \leqq t \int_0^N |\,q_m(u) - q(u)\,|\, du + |\,q_m(N-0) - q(N-0)\,| \\ &\qquad = o(1) \text{ as } m \to \infty. \end{aligned}$$

Combining these results we see that $H_m(t) \to H(t)$ as $m \to \infty$.

We propose to show that a necessary condition for L-effectiveness, if (2.2) is satisfied, is

$$\sum_{n=0}^{\infty} \max_{2^n \leqq t \leqq 2^{n+1}} |\,H(t)\,| < M. \tag{6.3}$$

If (6.3) is not satisfied then for every M, we can find an N such that

$$\sum_{n=0}^{N} |\,H(t_n)\,| > M,$$

where t_n is a point on $(2^n, 2^{n+1})$ where $H(t)$ attains its maximum absolute value. We choose $m > 2^N$ so that

$$| H_m(t_n) - H(t_n) | < \frac{M}{2N}, \qquad 0 \leqq n \leqq N.$$

Then

$$\sum_{n=0}^{[\log_2 m]} \max_{2^n \leqq t \leqq 2^{n+1}} | H_m(t) | > \sum_{n=0}^{N} | H_m(t_n) | - \frac{M}{2} > \frac{M}{2},$$

and hence H_m cannot be bounded. Therefore condition (6.3) is necessary for L-effectiveness.

After the same fashion it can be shown that, if (6.1) and (6.2) are satisfied,

$$\tilde{H}_m(t) \to \tilde{H}(t) = \int_0^\infty \cos ut\,dq(u) \text{ as } m \to \infty,$$

and, if (3.1) and (3.3) are satisfied, then a necessary condition for $\tilde{L}$-effectiveness is

$$\sum_{n=0}^{\infty} \max_{2^n \leqq t \leqq 2^{n+1}} | \tilde{H}(t) | < \infty. \tag{6.4}$$

7. **Methods of the closed cycle.** As an application of these results we shall consider the problem of methods of the closed cycle. For such methods $a_{mn} = a(n/m)$. The function $a(x)$ may be considered to be defined everywhere on $(0, \infty)$. Then (1.9) implies that

$$\int_0^\infty | da(x) | < A,$$

and we have, at every point of continuity of $a(u)$,

$$q_m(u) = a\left(\frac{[um]}{m}\right) - a(0) \to a(u) - a(0) \text{ as } m \to \infty.$$

so

$$\int_N^\infty | dq_m(u) | \leqq \int_N^\infty | da(u) | \to 0 \text{ as } N \to \infty,$$

so that the conditions of §6 are satisfied. Therefore, if (2.2) is satisfied a necessary condition for L-effectiveness is

$$\sum_{n=0}^{\infty} \max_{2^n \leqq t \leqq 2^{n+1}} | H(t) | < \infty, \tag{7.1}$$

where

$$H(t) = \int_0^\infty \sin ut\, da(u).$$

We propose to show that this condition is also sufficient. Since $H(t)$ is bounded and, by (2.2), $|H(t)| < Bt$, the function $(1/t)H(t) \subset L_2$ on $(0, \infty)$. Therefore it will have a Fourier transform contained in L_2, and since, if $a(x) \to 0$ as $x \to \infty$,

$$\frac{1}{t} H(t) = \int_0^\infty \cos ut\, a(u) du,$$

$(2/\pi)^{1/2} a(x)$ will be this transform. Condition (7.1) implies that $(1/t)H(t) \subset L$ and therefore

$$a(x) = \frac{2}{\pi} \int_0^\infty \cos xt \frac{1}{t} H(t) dt.$$

By the Fubini theorem, if $H(-t) = H(t)$, the function

$$\frac{1}{\pi} \int_{-\infty}^\infty f(x+t) \frac{1}{t} H(mt) dt$$

will be integrable on $(-\pi, \pi)$, since (7.1) implies that $(1/t)H(t) \subset L$ on $(0, \infty)$. Let us now consider its Fourier series. We have

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} \frac{1}{\pi} \int_{-\infty}^\infty f(x+t) \frac{1}{t} H(mt) dt dx \\ &= \frac{1}{\pi} \int_{-\infty}^\infty \frac{1}{t} H(mt) \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) e^{-inx} dx dt \\ &= \frac{1}{\pi} \int_{-\infty}^\infty \frac{1}{t} H(mt) f_n e^{int} dt = f_n a(|n|/m). \end{aligned}$$

Therefore

$$\frac{1}{\pi} \int_{-\infty}^\infty f(x+t) \frac{1}{t} H(mt) dt \sim \sum_{n=-\infty}^{\infty} a(|n|/m) \{ f_n e^{inx} + f_n e^{-inx} \},$$

or

$$T_m(L, x) = \frac{1}{\pi} \int_{-\infty}^\infty f(x+t) \frac{1}{t} H(mt) dt = \frac{1}{\pi} \int_0^\infty \phi(L, t) \frac{1}{t} H(mt) dt.$$

Now, if $x \subset E(L, f)$ and (2.2), (7.1) are satisfied,

$$\left| \frac{1}{\pi} \int_0^{1/m} \phi(L, t) \frac{1}{t} H(mt) dt \right| < \frac{Bm}{\pi} \int_0^{1/m} | \phi(L, t) | \, dt = o(1) \text{ as } m \to \infty,$$

and

$$\left| \int_a^\infty \phi(L, t) \frac{1}{t} H(mt) dt \right| = \left| \int_{ma}^\infty \phi\left(L, \frac{t}{m}\right) \frac{1}{t} H(t) dt \right|$$
$$\leqq \sum_{n=[\log_2 ma]}^\infty \max_{2^n \leqq t \leqq 2^{n+1}} | H(t) | \, 2^{-n} \int_0^{2^{n+1}} \left| \phi\left(L, \frac{t}{m}\right) \right| dt$$
$$= O\left(\sum_{n=[\log_2 ma]}^\infty \max_{2^n \leqq t \leqq 2^{n+1}} | H(t) | \right) = o(1) \text{ as } m \to \infty.$$

Finally

$$\left| \int_{1/m}^a \phi(L, t) \frac{1}{t} H(mt) dt \right| = \left| \int_1^{ma} \phi\left(L, \frac{t}{m}\right) \frac{1}{t} H(t) dt \right|$$
$$\leqq \sum_{n=0}^{[\log_2 ma]} \max_{2^n \leqq t \leqq 2^{n+1}} | H(t) | \, 2^{-n} \int_0^{2^{n+1}} \left| \phi\left(L, \frac{t}{m}\right) \right| dt$$
$$= o\left(\sum_{n=0}^\infty \max_{2^n \leqq t \leqq 2^{n+1}} | H(t) | \right) = o(1) \text{ uniformly in } m \text{ as } a \to 0.$$

From this it follows that (7.1) is necessary and sufficient for L-effectiveness, if (2.2) is satisfied. This result was conjectured by Paley in an unpublished paper but was not proved.

Similarly, if

$$\tilde{H}(t) = \int_0^\infty \cos ut \, da(u)$$

it can be shown that

$$T_m(\tilde{L}, x) = \frac{1}{\pi} \int_0^\infty \phi(\tilde{L}, t) \frac{1}{t} (-1 + \tilde{H}(mt)) dt,$$

and, if (3.1) and (3.3) are satisfied, a necessary and sufficient condition for $\tilde{L}$-effectiveness is

$$\sum_{n=0}^\infty \max_{2^n \leqq t \leqq 2^{n+1}} | \tilde{H}(t) | < \infty. \tag{7.2}$$

8. **Discussion.** From the preceding work there would seem to be apparent differences between F-effectiveness, L-effectiveness, and L'-effectiveness and also between $\tilde{F}$-effectiveness, $\tilde{L}$-effectiveness, and $\tilde{L}'$-effectiveness. Examples

have been given (see Paley, Randels, and Rosskopf [4] and Randels [5, Parts I and II]) showing that these differences are essential. It would be of interest to find relations between effectiveness for Fourier series and for conjugate series. One trivial result in this direction is possible, for we have seen that the L-effectiveness problem is not changed if $a_m = \lim_{n\to\infty} a_{mn} \neq 0$ while our discussion of $\tilde{L}$-effectiveness is only valid if $a_m = 0$. It is clear then that it is possible to construct methods of summability which are L-effective but not $\tilde{L}$-effective. It would be of much more interest to find out if this is still possible when $a_m = 0$.

It is natural to ask whether it would be possible to dispense with the conditions (2.2) and (3.1) and still get necessary and sufficient conditions. It appears to be quite definite that this method cannot accomplish this.

References

1. M. Fekete, *Über Faktorenfolgen welche die "Klasse" einer Fourierschen Reihe unverändert lassen*, Acta Szeged, Universität Franz Joseph, vol. 1 (1923), pp. 148–166.

2. E. Hille and J. D. Tamarkin, *On the summability of Fourier series*, I, these Transactions, vol. 34 (1932), pp. 757–783.

3. E. Hille and J. D. Tamarkin, *On the summability of Fourier series*, fifth note, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 369–372.

4. R. E. A. C. Paley, W. C. Randels, and M. F. Rosskopf, *On the summation of Fourier series*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 69–74.

5. W. C. Randels, *Three examples in the theory of Fourier series*, Annals of Mathematics, (2), vol. 36 (1935), pp. 838–858.

Institute for Advanced Study,
Princeton, N. J.

CONVERGENCE IN VARIATION AND RELATED TOPICS†

BY

ANTHONY P. MORSE

1. **Introduction.** In recent papers‡ by Adams and Clarkson and by Adams and Lewy the notions of convergence in variation and convergence in length have been examined. In AC it has been shown that if a sequence converges in variation and satisfies certain further restrictions which are clearly needed, the sequence of reciprocals converges in variation. The central purpose of the present paper is to determine so far as we are able the transformations which when applied to sequences of functions, preserve various types of convergence, such as convergence in variation or length and other types which we shall introduce. This paper also leads us to certain generalizations of results in AC and AL, such as Theorem 5.4 wherein convergence in length is seen to be invariant under addition and multiplication when *only one* of the limit functions is absolutely continuous.

In §2 we assemble certain preliminary definitions, notations, and conventions. §3 is devoted to preliminary theorems and lemmas, among them being Theorems 3.1 and 3.2 which might be of interest in themselves, their full power, in fact, not being used in this paper. Theorem 3.2 is a substitution theorem for Lebesgue integrals which is more general than other theorems of this type known to us in literature. Certain results in §3 are, however, obvious analogues of results in AC. Transforms of sequences are discussed in §4, wherein Theorems 4.1 and 4.2 form the kernel of the paper. The remainder of the paper consists largely of various applications of these two theorems, convergence in length being discussed in §5 together with convergence almost in the mean, uniform convergence in length in §6, and strong convergence in §7. Certain miscellaneous applications are made in §8; these include Theorem 8.1 which points out a necessary and sufficient condition for convergence in the mean, and Theorems 8.3 and 8.4 which are generalizations of a theorem of Plessner.

2. **Notation; preliminary definitions and conventions.** In this paper we shall consistently use x, y, t to denote real numbers or variables, and use the

---

† Presented to the Society, October 31, 1936; received by the editors May 22, 1936.

‡ Adams and Clarkson, *On convergence in variation*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 413–417. Adams and Lewy, *On convergence in length*, Duke Mathematical Journal, vol. 1 (1935), pp. 19–26. Hereinafter these papers will be referred to as AC and AL, respectively.

letters X, Y, ξ, ψ, Ψ, η, u, v, U, V with or without subscripts to denote real-valued functions. All other functions are to be regarded as complex functions of a real variable unless the contrary is expressly stated. We shall also employ a and b with or without subscripts to designate real numbers with $a<b$.

If Q is a condition involving x, then $\underset{x}{E}[Q]$ is a set defined as follows: A point x belongs to $\underset{x}{E}[Q]$ if x satisfies the condition Q. We use the notation $[x_1, x_2]$ to denote a closed interval.

If f is a function and if $[a, b]$ is included in its domain, then the symbol $T_a^b(f)$ (read the total variation from a to b of f) will be used to denote the least upper bound (finite or infinite) of numbers of the form

$$\sum_{j=1}^{k} | f(t_j) - f(t_{j-1}) |,$$

where $a=t_0<t_1<t_2< \cdots <t_k=b$. We define $T_b^a(f)=-T_a^b(f)$ and $T_a^a(f)=0$. If $T_a^b(f)<\infty$, then f is said to be of b.v. (bounded variation) on $[a, b]$. Since it will sometimes be necessary to display the variable with respect to which the total variation is taken, we employ the notation $T_{t=a}^{b}f(t)$ as an alternate for $T_a^b(f)$.

We shall use a.c. as an abbreviation for absolute continuity and employ p.p. to denote almost everywhere (presque partout), and designate the outer measure of a set R by $|R|$. It will also be convenient to refer to Euclidean space of n dimensions as simply n-space. Furthermore, a function will be said to be increasing on a set if it is strictly increasing there, a function will be said to be monotone on a set if it is either non-increasing or non-decreasing there.

The following convention, will be used throughout the paper. If f is defined on $[a, b]$, then the function f' is defined on $[a, b]$ by the following relations:

$f'(t)=$ the derivative of f at t wherever it exists finite,

$f'(t)=0$ for all other t on $[a, b]$.

We also agree: If $[a, b]$ is the domain of f and if $\lim_{h\to 0+}f(t+h)$ exists for $a\leqq t<b$ and if $\lim_{h\to 0+}f(t-h)$ exists for $a<t\leqq b$, then $f(t+)$ and $f(t-)$ are defined for t on $[a, b]$ as follows:

$$f(a-) = f(a); \quad f(b+) = f(b); \quad f(t+) = \lim_{h\to 0+} f(t+h), \quad a \leqq t < b;$$

$$f(t-) = \lim_{h\to 0+} f(t-h), \quad a < t \leqq b.$$

We shall denote by CR the set of all finite-valued complex functions whose domain is $[0, 1]$. If f and g are any points in CR and if α is any complex number, then by $f+g$ is meant that point G in CR such that $G(t)=f(t)+g(t)$ for t on $[0, 1]$, by $f\cdot g$ is meant that point G in CR such that $G(t)=f(t)\cdot g(t)$ for t on

[0, 1], by αf is meant that point G in CR such that $G(t)=(\alpha)\cdot f(t)$ for t on [0, 1]; and, provided f does not vanish on [0, 1], by α/f is meant that point G in CR such that $G(t)=\alpha/f(t)$ for t on [0, 1]. We denote by I and θ the elements of CR defined respectively by

$$I(t) = t, \qquad \theta(t) = 0; \qquad 0 \leqq t \leqq 1.$$

If f is in CR, then $\|f\|$, read norm of f, is defined as $|f(0)|+T_0^1(f)$. The space BV is a subspace of CR defined by $BV=(CR)\cdot \underset{f}{E}[\|f\|<\infty]$ and the subspace RBV is defined by $RBV=(BV)\cdot \underset{f}{E}$ [f is real].

Totally distinct from these spaces and used merely for convenience is the space CC defined as the space of continuous functions on the finite complex plane to the finite complex plane.

In concluding this section we lay down the following definitions.

DEFINITION 2.1. If f is a point in CR and ϕ is a function whose domain includes the range of f and whose range is included in the set of finite complex numbers (finite complex plane), then $\phi:f$ is defined to be that point G in CR for which $G(t)=\phi\{f(t)\}$, $0\leqq t\leqq 1$.

DEFINITION 2.2. If Y is a real point in CR and u is a function on a part of two-space to one-space and if $u(t, Y(t))$ is defined for t on [0, 1], then by $(u|Y)$ is meant the real point Ψ in CR such that $\Psi(t)=u(t, Y(t))$, $0\leqq t\leqq 1$.

DEFINITION 2.3. *Convergence in variation.* By $f_n-v\rightarrow f_0$, read f_n converges in variation to f_0, is meant this: f_n is in BV for $n=0, 1, 2, \cdots$; $f_n(t)\rightarrow f_0(t)$ for t on [0, 1]; $\|f_n\|\rightarrow\|f_0\|$.

DEFINITION 2.4. *Uniform convergence in variation.* By $f_n-uv\rightarrow f_0$ is meant this: $f_n-v\rightarrow f_0$ with $f_n(t)\rightarrow f_0(t)$ uniformly for t on [0, 1].

DEFINITION 2.5. *Convergence in length.* By $Y_n-l\rightarrow Y_0$ is meant this: Y_n is in RBV for $n=0, 1, 2, \cdots$; $(I+iY_n)-v\rightarrow(I+iY_0)$.

DEFINITION 2.6. *Uniform convergence in length.* By $Y_n-ul\rightarrow Y$ is meant this: $Y_n-l\rightarrow Y_0$; $Y_n(t)\rightarrow Y_0(t)$ uniformly for t on [0, 1].

DEFINITION 2.7. *Strong convergence.* By $f_n-s\rightarrow f_0$ is meant this: f_n is in BV for $n=0, 1, 2, \cdots$; $\|f_n-f_0\|\rightarrow 0$.

DEFINITION 2.8. If α is a point in n_1-space with $\alpha=(\alpha_1, \alpha_2, \cdots, \alpha_{n_1})$ and β a point in n_2-space with $\beta=(\beta_1, \beta_2, \cdots, \beta_{n_2})$, then $\alpha\circ\beta$ is the point $(\alpha_1, \alpha_2, \cdots, \alpha_{n_1}, \beta_1, \beta_2, \cdots, \beta_{n_2})$ in (n_1+n_2)-space.

3. **Preliminary results.** In this section certain preliminary results will be actually proved while others which are quite simple or well known will simply be stated both for completeness and for use later.

From the definition of total variation it is clear that if $a\leqq c\leqq b$ with f defined on $[a, b]$, then $T_a^b(f)=T_a^c(f)+T_c^b(f)$. Another corollary is the fol-

lowing semi-continuity property: the relation $f_n(t)\to f_0(t)$ for t on $[a, b]$ implies $\lim\inf_{n\to\infty} T_a^b(f_n) \geqq T_a^b(f_0)$. It is likewise easily verified that the relation $f_n - v\to f_0$ with f_0 continuous implies $f_n - uv\to f_0$ (see AC, Theorem 2 and corollary to Theorem 4).

Of considerable use is the following

LEMMA 3.1. *The relations $f_n - v\to f_0$ and $f_n - uv\to f_0$ imply respectively the relations*† *$S_n(t)\to S_0(t)$ for t on $[0, 1]$; $S_n(t)\to S_0(t)$, uniformly for t on $[0, 1]$ where $S_n(t) = T_0^t(f_n)$ for t on $[0, 1]$.*

Several properties of the norm in CR are now set forth in the following lemma.

LEMMA 3.2. *If f and g are in CR and α is a complex number, then*

$$\|f + g\| \leqq \|f\| + \|g\|, \quad \|\alpha f\| = |\alpha| \cdot \|f\|, \quad \|fg\| \leqq \|f\|\cdot\|g\|.$$

The first two relations are quite simple. We sketch a proof of the last which is seen to reduce to proving

$$\left(\sum_{n=0}^{N} | a_n - a_{n-1} |\right)\cdot\left(\sum_{n=0}^{N} | b_n - b_{n-1} |\right) \geqq \sum_{n=0}^{N} | a_n b_n - a_{n-1}b_{n-1} |,$$

where a_n and b_n are complex numbers ($n = 0, 1, \cdots, N$) and $|a_{-1}| + |b_{-1}| = 0$. Suppose the above relation, which is obviously true if N is replaced by 0, to be true for N replaced by an integer k with $0 \leqq k < N$. Then it follows that

$$\begin{aligned}
&\left(| a_{k+1} - a_k | + \sum_{n=0}^{k} | a_n - a_{n-1} |\right)\cdot\left(| b_{k+1} - b_k | + \sum_{n=0}^{k} | b_n - b_{n-1} |\right) \\
&\quad = | a_{k+1} - a_k | \cdot | b_{k+1} - b_k | + | a_{k+1} - a_k | \cdot \sum_{n=0}^{k} | b_n - b_{n-1} | \\
&\qquad + | b_{k+1} - b_k | \cdot \sum_{n=0}^{k} | a_n - a_{n-1} | + \left(\sum_{n=0}^{k} | a_n - a_{n-1} |\right)\cdot\left(\sum_{n=0}^{k} | b_n - b_{n-1} |\right) \\
&\quad \geqq | (a_{k+1} - a_k)(b_{k+1} - b_k) + (a_{k+1} - a_k)b_k + (b_{k+1} - b_k)a_k | \\
&\qquad\qquad + \sum_{n=0}^{k} | a_n b_n - a_{n-1}b_{n-1} | = \sum_{n=0}^{k+1} | a_n b_n - a_{n-1}b_{n-1} |.
\end{aligned}$$

This induction completes the proof.

The following two lemmas are particular cases of Minkowski's inequality.

† AC, corollaries to Theorems 2 and 5. The proof given there holds equally well for the functions considered here.

LEMMA 3.3. *If* $0 \leqq a_j \leqq b_j$ *(for* $j=1, 2, \cdots, k$*), then*

$$\sum_{j=1}^{k} (b_j^2 - a_j^2)^{1/2} \leqq \left[\left(\sum_{j=1}^{k} b_j\right)^2 - \left(\sum_{j=1}^{k} a_j\right)^2\right]^{1/2}.$$

LEMMA 3.4. *If X and Y are summable on* $[a, b]$, *then*

$$\int_a^b \{ | X(t) | + | Y(t) | \} dt \geqq \int_a^b (\{ X(t)\}^2 + \{Y(t)\}^2)^{1/2} dt$$
$$\geqq \left\{\left(\int_a^b | X(t) | \, dt\right)^2 + \left(\int_a^b | Y(t) | \, dt\right)^2\right\}^{1/2}.$$

LEMMA 3.5. *If X is a function defined on* $[a, b]$, *then* X' *is measurable on* $[a, b]$.

This is a corollary of a theorem found in Saks, *Théorie de l'Intégrale* (Chapter 3, p. 47, Theorem 1).

LEMMA 3.6. *Let* $[a, b]$ *be the domain of the function X and denote* $\underset{t}{E}[|X'(t)| > 0]$ *by P. If D is a set of measure* 0, *then the set* $P \cdot \underset{t}{E}[X(t) \varepsilon D]$ *is likewise of measure* 0.

Let $Q = \underset{t}{E}[X(t)\varepsilon D]$ and suppose $|PQ| > 0$. Define $P_1 = \underset{t}{E}[X'(t) > 0]$. There is no loss of generality in assuming $|P_1Q| > 0$. P_1 being measurable, it is easily established that there exist positive numbers ϵ_1, ϵ_2, and a closed set C such that

$$C \subset P_1, \quad | P_1 - C | < | P_1Q |, \quad \frac{X(t_2) - X(t_1)}{t_2 - t_1} \geqq \epsilon_1$$

if $t_1 \varepsilon C$, $t_2 \varepsilon C$, and $0 < |t_2 - t_1| < \epsilon_2$. Thus $|CQ| > 0$, so that there exists an interval $[a_0, b_0]$ of length $< \epsilon_2$ for which $|CQ[a_0, b_0]| > 0$. Defining $C^* = C \cdot [a_0, b_0]$ and $Q^* = C^* \cdot Q$ we observe $|Q^*| > 0$ and C^* is closed, so that if X_1 is defined on C^* so as to coincide with X and defined on the remainder of $[a_0, b_0]$, which is made up of a set of non-overlapping intervals, by linear interpolation then it appears that X_1 is continuous and increasing on $[a_0, b_0]$ with

$$| X_1(t_2) - X(t_1) | \geqq \epsilon_1 | t_2 - t_1 |, \ (a_0 \leqq t_1, t_2 \leqq b_0); \ X_1(t) = X(t), \ (t \varepsilon C^*).$$

The continuity of X_1 follows from the existence of the derivative of X at all points of C^*. Let τ be defined on the interval $[X_1(a_0), X_1(b_0)]$ by the relation $\tau\{X_1(t)\} = t$ $(a_0 \leqq t \leqq b_0)$. Clearly τ satisfies a Lipschitz condition on $[X_1(a_0), X_1(b_0)]$, and thus transforms the set $D \cdot [X_1(a_0), X_1(b_0)]$ which is of measure 0, into a set D^* of measure 0. But

$$Q^* = C^*Q = C^*\underset{t}{E}[X(t)\varepsilon D] = C^*\underset{t}{E}[X_1(t)\varepsilon D] = C^*D^*.$$

Hence the contradiction $|Q^*| = 0$.

COROLLARY 3.1. *Let $[a, b]$ be the domain of X and let $[\alpha, \beta]$ include its range. If ψ_1 and ψ_2 are finite-valued functions defined on $[\alpha, \beta]$ and equal p.p. there, then for almost all t on $[a, b]$, $\psi_1\{X(t)\}X'(t) = \psi_2\{X(t)\}X'(t)$.*

THEOREM 3.1. *Let $[a, b]$ be the domain of the function X; let $[\alpha, \beta]$ include its range; let Ψ be a.c. on $[\alpha, \beta]$ and denote $\Psi\{X(t)\}$ by $G(t)$ for t on $[a, b]$. If X has a derivative p.p. on $[a, b]$, then*

$$G'(t) = \Psi'\{X(t)\}X'(t)$$

for almost all t on $[a, b]$.

Let it first be noted that the theorem does not answer the question as to whether or not G is differentiable p.p.† We now proceed with the proof of the theorem.

Let $R = \underset{t}{E}[X'(t) = 0]$. If $|R| = 0$ the next paragraph by itself furnishes a proof. However, in case $|R| > 0$, it may be noted first that corresponding to $\epsilon(>0)$ there exist positive numbers M, ϵ_1 and a closed subset C, of R, such that $|R - C| < \epsilon$ with $|X(t_2) - X(t_1)| \leqq M|t_2 - t_1|$ if $t_1 \epsilon C$, $t_2 \epsilon C$, and $|t_2 - t_1| < \epsilon_1$. Likewise readily verified is the existence of a function X_1 satisfying a Lipschitz condition on $[a, b]$ (the constant involved may be $> M$), and in addition fulfilling: $X_1(t) = X(t)$ for $t \epsilon C$. Now X_1 and X both transform C into a measurable (closed) set C' with

$$|C'| \leqq \int_C |X_1'(t)| \, dt = \int_C |X'(t)| \, dt = 0.$$

But Ψ being a.c. on $[\alpha, \beta]$, transforms C' into a set C'' likewise of measure 0. Hence G transforms C into C'', a set of measure 0, so that $G'(t) = 0$ for almost all t in C; for supposing the contrary leads immediately to a contradiction of Lemma 3.6. Hence, since ϵ was arbitrary, we conclude $G'(t) = 0$ for almost all t in R. Thus $G'(t) = \Psi'\{X(t)\}X'(t)$ for almost all t in R.

However, upon denoting $\underset{t}{E}[|X'(t)| > 0]$ by P, it is seen from Lemma 3.6 that

$$\lim_{h \to 0} \frac{\Psi\{X(t+h)\} - \Psi\{X(t)\}}{X(t+h) - X(t)} = \Psi'\{X(t)\}$$

† In fact, from the work of N. Bary, Mathematische Annalen, vol. 103 (1930), p. 611, a definite answer to this question can be given. Let F be a continuous function nowhere differentiable on $[0, 1]$. There then exist functions G_1, G_2, G_3 and a.c. functions Ψ_1, Ψ_2, Ψ_3, ϕ_1, ϕ_2, ϕ_3 such that

$$F(t) = G_1(t) + G_2(t) + G_3(t), \quad G_j(t) = \Psi_j\{\phi_j(t)\} \quad (j = 1, 2, 3; 0 \leqq t \leqq 1).$$

Thus at least one of the functions G_1, G_2, G_3 fails to have a derivative on a set of measure >0.

for almost all t in P. Hence $G'(t)=\Psi'\{X(t)\}X'(t)$ for almost all t on $[a, b]$.†

COROLLARY 3.2. *Let $[a, b]$ be the domain of X whereon it is a.c. and let $[\alpha, \beta]$ include its range. If Ψ satisfies a Lipschitz condition on $[\alpha, \beta]$ or if Ψ is a.c. on $[\alpha, \beta]$ with X monotone on $[a, b]$, then the function G defined by $G(t)=\Psi\{X(t)\}$ $(a\leqq t\leqq b)$, is a.c. on $[a, b]$ with $G'(t)=\Psi'\{X(t)\}\ X'(t)$ for almost all t on $[a, b]$.*

COROLLARY 3.3. *Let $[a, b]$ be the domain of X; let $[\alpha, \beta]$ include its range; let Ψ be a.c. on $[\alpha, \beta]$; and denote $\Psi\{X(t)\}$ by $G(t)$ for t on $[a, b]$. If X has a vanishing derivative p.p. on $[a, b]$, then a necessary and sufficient condition that G is a.c. on $[a, b]$ is that Ψ is constant on the range of X.*

Now combining Theorem 3.1 with Corollary 3.1 we obtain

THEOREM 3.2. *Let X be defined on $[a, b]$ and differentiable p.p. there; let $[\alpha, \beta]$ include the range of X; let ψ be defined, finite-valued, and summable on $[\alpha, \beta]$ and denote $\int_a^x\psi(s)ds$ by $\Psi(x)$ for x on $[\alpha, \beta]$; finally let $G(t)=\Psi\{X(t)\}$ for t on $[a, b]$. Under these circumstances*

$$\int_{X(a)}^{X(t)}\psi(x)dx=\int_a^t\psi\{X(s)\}X'(s)ds \qquad (a\leqq t\leqq b)$$

if and only if G is a.c. on $[a, b]$.‡

LEMMA 3.7. *If f is defined on $[a, b]$ then*

$$T_a^b(f)\geqq\int_a^b|f'(t)|\,dt,$$

the sign of equality holding if f is a.c. on $[a, b]$.

---

† A slight modification in proof establishes:

If Ψ satisfies Lusin's condition N on $[\alpha, \beta]$ (i.e., transforms sets of measure 0 into sets of measure 0) and if X is continuous on $[a, b]$ in addition to being differentiable p.p. there, then $G'(t)=\Psi'\{X(t)\}X'(t)$ for almost all t on $[a, b]$.

It may also be noted that a slight simplification in proof may be achieved by use of a theorem of N. Bary, loc. cit., p. 190.

‡ This theorem is a generalization of results previously obtained by de la Vallée Poussin, these Transactions, vol. 16 (1915), p. 466, and by Fichtenholz, Bulletin de l'Académie Royale de Belgique, Classe des Sciences, ser. 5, vol. 8 (1922), p. 441. In the notation of the present theorem it was shown by de la Vallée Poussin that absolute continuity of X implies the equivalence of the following three statements:

(i) $\int_{X(a)}^{X(b)}\psi(x)dx=\int_a^b\psi\{X(s)\}X'(s)ds$;

(ii) G is a.c.;

(iii) $\int_a^b\psi\{X(s)\}X'(s)ds$ exists.

On the other hand Fichtenholz showed that absolute continuity of G together with monotonicity and continuity of X implies (i).

LEMMA 3.8. *Let σ be monotone and continuous on $[a, b]$. If f is defined on $[\sigma(a), \sigma(b)]$, then*

$$T_{t=a}^{b} f\{\sigma(t)\} = | T_{\sigma(a)}^{\sigma(b)}(f) |.$$

LEMMA 3.9. *If f is a.c. on $[a, b]$ with $f'(t) = \mu$ (constant) for almost all t on $[a, b]$, then*

$$\int_a^b | f'(t) - \Delta\delta^{-1} | \, dt \leqq (\mu^2\delta^2 - | \Delta |^2)^{1/2} + 2(\mu\delta - | \Delta |),$$

where $\delta = b - a$ and $\Delta = f(b) - f(a)$.

The lemma is clearly true in case $\Delta = 0$. In the alternative case we define

$$F(t) = | \Delta | (f(t) - f(a))\Delta^{-1} = \xi(t) + i\eta(t) \qquad (a \leqq t \leqq b).$$

Now (Lemma 3.4)

$$\begin{aligned}
\mu\delta &= \int_a^b | f'(t) | \, dt = \int_a^b | F'(t) | \, dt = \int_a^b (\{\xi'(t)\}^2 + \{\eta'(t)\}^2)^{1/2} dt \\
&\geqq \left(\left\{ \int_a^b | \xi'(t) | \, dt \right\}^2 + \left\{ \int_a^b | \eta'(t) | \, dt \right\}^2 \right)^{1/2} \\
&\geqq \left(\left\{ \int_a^b \xi'(t) dt \right\}^2 + \left\{ \int_a^b | \eta'(t) | \, dt \right\}^2 \right)^{1/2} \\
&= \left(| \Delta |^2 - \left\{ \int_a^b | \eta'(t) | \, dt \right\}^2 \right)^{1/2}
\end{aligned}$$

whence $(\mu^2\delta^2 - |\Delta|^2) \geqq \{\int_a^b |\eta'(t)| dt\}^2$, so that from the relations

$$\int_a^b \left| | \Delta | f'(t)\Delta^{-1} - \xi'(t) \right| dt = \int_a^b | \eta'(t) | \, dt,$$

$$\int_a^b | \xi'(t) - \mu | \, dt = \int_a^b \{\mu - \xi'(t)\} dt = \mu\delta - | \Delta |,$$

$$\int_a^b \left| \mu - | \Delta | \delta^{-1} \right| dt = \mu\delta - | \Delta |,$$

follows the relation

$$\begin{aligned}
\int_a^b | f'(t) - \Delta\delta^{-1} | \, dt &= \int_a^b \left| | \Delta | f'(t)\Delta^{-1} - | \Delta | \delta^{-1} \right| dt \\
&\leqq (\mu^2\delta^2 - | \Delta |^2)^{1/2} + 2(\mu\delta - | \Delta |).
\end{aligned}$$

THEOREM 3.3. *If f_n is an a.c. function in BV with $|f_n'(t)| = \mu_n$ for almost all t on $[0, 1]$ $(n=0, 1, 2, \cdots)$, then the relation $f_n - v \to f_0$ implies $\|f_n - f_0\| \to 0$.*

From

$$\mu_n = \int_0^1 |f_n'(t)|\,dt = T_0^1(f_n) \qquad (n = 0, 1, 2, \cdots)$$

we conclude $\mu_n \to \mu_0$. Let ϵ be any positive number. There exists a partition $(0 = t_0 < t_1 < t_2 < \cdots < t_k = 1)$ such that $0 \leqq \mu_0 - \sum_{j=1}^k |f_0(t_j) - f_0(t_{j-1})| < \epsilon$. Letting $\Delta_{n,j} = f_n(t_j) - f_n(t_{j-1})$ and $\delta_j = t_j - t_{j-1}$, $(n=0, 1, 2, \cdots; j = 1, 2, \cdots, k)$, we have the obvious relation

$$\sum_{j=1}^k \int_{t_{j-1}}^{t_j} |\Delta_{n,j}\delta_j^{-1} - \Delta_{0,j}\delta_j^{-1}|\,dt = \sum_{j=1}^k |\Delta_{n,j} - \Delta_{0,j}| \qquad (n = 0, 1, 2, \cdots).$$

Use of Lemma 3.9 and Lemma 3.3 yields (for $n = 0, 1, 2, \cdots$)

$$\sum_{j=1}^k \int_{t_{j-1}}^{t_j} |f_n'(t) - \Delta_{n,j}\delta_j^{-1}|\,dt \leqq \sum_{j=1}^k (\mu_n^2\delta_j^2 - |\Delta_{n,j}|^2)^{1/2}$$

$$+ \sum_{j=1}^k 2(\mu_n\delta_j - |\Delta_{n,j}|) \leqq K_n,$$

where

$$K_n = \left\{\mu_n^2 - \left(\sum_{j=1}^k |\Delta_{n,j}|\right)^2\right\}^{1/2} + 2\left\{\mu_n - \sum_{j=1}^k |\Delta_{n,j}|\right\}.$$

Combining these last two relations with the relation obtained by setting $n = 0$ in the last relation we conclude

$$\int_0^1 |f_n'(t) - f_0'(t)|\,dt = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} |f_n'(t) - f_0'(t)|\,dt$$

$$\leqq K_n + K_0 + \sum_{j=1}^k |\Delta_{n,j} - \Delta_{0,j}| \qquad (n = 0, 1, 2, \cdots).$$

Hence

$$\limsup_{n\to\infty} \int_0^1 |f_n'(t) - f_0'(t)|\,dt$$

$$\leqq 2K_0 = 2\left\{\left(\mu_0 - \sum_{j=1}^k |\Delta_{0,j}|\right)\left(\mu_0 + \sum_{j=1}^k |\Delta_{0,j}|\right)\right\}^{1/2} + 4\left(\mu_0 - \sum_{j=1}^k |\Delta_{0,j}|\right)$$

$$\leqq 2(\epsilon \cdot 2\mu_0)^{1/2} + 4\epsilon,$$

the theorem following from the arbitrariness of ϵ.

4. **Transforms of sequences in BV.** We prove the following lemma.

LEMMA 4.1. *If* $\{X_n\}$ *is a sequence of functions in RBV satisfying the condition*

$$|X_n(t_2) - X_n(t_1)| \leqq M|t_2 - t_1| \qquad (0 \leqq t_1, t_2 \leqq 1;\ n = 0, 1, 2, \cdots),$$

and if Ψ *is a real function which satisfies a Lipschitz condition on every finite interval, then the relation* $\|X_n - X_0\| \to 0$ *implies the relation* $\|\Psi:X_n - \Psi:X_0\| \to 0$.

Let $[a, b]$ be such that $X_n(t)$ is in $[a, b]$ for $0 \leqq t \leqq 1$; $n = 0, 1, 2, \cdots$. Let $M_1(>0)$ dominate $\Psi'(x)$ for $a \leqq x \leqq b$; let $P_n = \underset{t}{E}[X_n'(t) > 0]$, $R_n = \underset{t}{E}[X_n'(t) = 0]$, $N_n = \underset{t}{E}[X_n'(t) < 0]$ for $n = 0, 1, 2, \cdots$ and denote by P^*, R^*, N^* those points on $[0, 1]$ at which the metric density of P_0, R_0, N_0 respectively is 1; let $\{\psi_p\}$ be a sequence of functions continuous on $[a, b]$ and dominated there by M_1, such that $\int_a^b H_p(x)dx \to 0$, where $H_p(x) = |\Psi'(x) - \psi_p(x)|$ for x on $[a, b]$. Since $\Psi:X_n$ satisfies a Lipschitz condition, it becomes clear in the light of Corollary 3.2 that the truth of the theorem is equivalent to showing $F_n(t) \to 0$ for t on $[0, 1]$, where F_n is defined by

$$F_n(t) = \int_0^t |\Psi'\{X_n(s)\}X_n'(s) - \Psi'\{X_0(s)\}X_0'(s)|\,ds.$$

To establish this relation we shall prove the following: If $\{F_n^*\}$ is any subsequence of $\{F_n\}$, then a subsequence $\{G_n\}$ of $\{F_n^*\}$ exists such that $G_n(t) \to 0$ for t on $[0, 1]$.

Since the sequence $\{F_n\}$ is comprised of non-decreasing functions which uniformly satisfy a Lipschitz condition, it appears as a corollary of Helly's theorem that there exists a subsequence $\{G_n\}$ of $\{F_n^*\}$ and a function G_0 satisfying a Lipschitz condition such that $G_n(t) \to G_0(t)$ for t on $[0, 1]$.

Now let t_0 be any point of P^* and denote $[t_0, s]$ by Q_s for $0 \leqq s \leqq 1$. From Theorem 3.2 follows the relation

$$\left|\int_{t_0}^{s} |\Psi'\{X_n(t)\}X_n'(t) - \psi_p\{X_n(t)\}X_n'(t)|\,dt\right| = \left|\int_{t_0}^{s} H_p\{X_n(t)\}\,|X_n'(t)|\,dt\right|$$

$$= \left|\int_{t_0}^{s} H_p\{X_n(t)\}X_n'(t)dt\right| + 2\int_{Q_sN_n} H_p\{X_n(t)\}\,|X_n'(t)|\,dt$$

$$\leqq \int_a^b H_p(x)dx + 4M_1M\,|Q_sN_n|$$

$$\leqq \int_a^b H_p(x)dx + 4M_1M\{|Q_sP_0N_n| + |Q_s(N_0 + R_0)|\}$$

$$(n = 0, 1, 2, \cdots;\ p = 1, 2, 3, \cdots;\ 0 \leqq s \leqq 1).$$

Thus

$$\left|\int_{t_0}^{s} \left| \psi_p\{X_0(t)\}X_0'(t) - \Psi'\{X_0(t)\}X_0'(t) \right| dt\right|$$
$$\leqq \int_a^b H_p(x)dx + 4M_1M \left| Q_s(N_0 + R_0) \right|$$

for $p=1, 2, 3,\cdots; 0 \leqq s \leqq 1$; furthermore, since X_n' converges in measure to X_0', it may easily be seen that $|Q_sP_0N_n| \to 0$ as $n \to \infty$ and hence that

$$\limsup_{n\to\infty} \left|\int_{t_0}^{s} \left| \Psi'\{X_n(t)\}X_n'(t) - \psi_p\{X_n(t)\}X_n'(t) \right| dt \right|$$
$$\leqq \int_a^b H_p(x)dx + 4M_1M \left| Q_s(N_0 + R_0) \right|$$

for $p=1, 2, 3, \cdots; 0 \leqq s \leqq 1$. By combining these two relations with the obvious relation

$$\limsup_{n\to\infty} \left|\int_{t_0}^{s} \left| \psi_p\{X_n(t)\}X_n'(t) - \psi_p\{X_0(t)\}X_0'(t) \right| dt\right| = 0$$
$$(p = 1, 2, 3, \cdots; 0 \leqq s \leqq 1),$$

we obtain

$$\begin{aligned}\limsup_{n\to\infty} &\left| F_n(s) - F_n(t_0) \right| \\ &= \limsup_{n\to\infty} \left|\int_{t_0}^{s} \Psi'\{X_n(t)\}X_n'(t) - \Psi\{X_0'(t)\}X_0'(t)\,dt\right| \\ &\leqq 2\int_a^b H_p(x)dx + 8M_1M \left| Q_s(N_0 + R_0) \right| \quad (p = 1, 2, 3, \cdots; 0 \leqq s \leqq 1)\end{aligned}$$

so that upon letting $p \to \infty$ we conclude

$$\begin{aligned}\left| G_0(s) - G_0(t_0) \right| &= \lim_{n\to\infty} \left| G_n(s) - G_n(t_0) \right| \\ &\leqq \limsup_{n\to\infty} \left| F_n(s) - F_n(t_0) \right| \leqq 8M_1M \left| Q_s(N_0 + R_0) \right|\end{aligned}$$

for s on $[0, 1]$. Thus, since t_0 is a point at which the metric density of P_0 is 1,

$$\lim_{s\to t_0} \frac{\left| Q_s(N_0 + R_0) \right|}{\left| s - t_0 \right|} = 0$$

which implies $G_0'(t_0) = 0$.

If t_0 is a point of N^*, a similar proof establishes $G_0'(t_0) = 0$; if $t_0 \epsilon R^*$, then the relation $G_0'(t_0) = 0$ is a consequence of the easily proved inequality

$$\limsup_{n\to\infty} \left| F_n(s) - F_n(t_0) \right| \leqq 2M_1M \left| Q_s(N_0 + P_0) \right| \quad (0 \leqq s \leqq 1).$$

Thus $G_0'(t)=0$ for almost all t on $[0, 1]$, and hence $G_0(t)=0$ for t on $[0, 1]$. This completes the proof.

The following is now readily established.

LEMMA 4.2. *If $\{X_n\}$ is a sequence of monotone functions in RBV satisfying the condition*

$$|X_n(t_2) - X_n(t_1)| \leqq M |t_2 - t_1| \qquad (0 \leqq t_1, t_2 \leqq 1; n = 0, 1, 2, \cdots),$$

and if Ψ is a real function which is a.c. on every finite interval, then the relation $\|X_n - X_0\| \to 0$ implies $\|\Psi:X_n - \Psi:X_0\| \to 0$.

To show this, approximate Ψ' in the mean with bounded measurable functions and apply Lemma 4.1 and Theorem 3.2. It is of some interest to note that some of the functions which comprise $\{X_n\}$ may be increasing while others are decreasing.

DEFINITION 4.1. Let u be a function on two-space to one-space. If there is a function A on one-space to n_1-space, a function B on one-space to n_2-space, and a function U on (n_1+n_2)-space to one-space such that

(i) $A(x) = (A_1(x), A_2(x), \cdots, A_{n_1}(x))$, $B(y) = (B_1(y), B_2(y), \cdots, B_{n_2}(y))$ $(-\infty < x, y < \infty)$;

(ii) all first partial derivatives of $U(\alpha_1, \alpha_2, \cdots, \alpha_{n_1}, \beta_1, \beta_2, \cdots, \beta_{n_2})$ with respect to these arguments are continuous functions on (n_1+n_2)-space;

(iii) $u(x, y) = U(A(x) \circ B(y)) = U(A_1(x), \cdots, A_{n_1}(x), B_1(y), \cdots, B_{n_2}(y))$ for $-\infty < x, y < \infty$;

(iv) on every finite interval either

A and B are a.c., or

A satisfies a Lipschitz† condition and B is a.c., or

A is a.c. and B satisfies a Lipschitz condition, or

A and B satisfy a Lipschitz condition;

then u is said to be respectively either $\mathfrak{K}$ or $\mathfrak{K}_1$ or $\mathfrak{K}_2$ or $\mathfrak{K}_{12}$.

DEFINITION 4.2. Let ϕ be a function in CC with $\phi(x+iy) = u(x, y) + iv(x, y)$ for $-\infty < x, y < \infty$; let f be a point in BV with $f = X + iY$. If the functions u and v are $\mathfrak{K}$ with X and Y monotone, or $\mathfrak{K}_1$ with Y monotone, or $\mathfrak{K}_2$ with X monotone, or $\mathfrak{K}_{12}$, then ϕ is said to be *applicable* to f.

DEFINITION 4.3. Let Y be a point in RBV and let u be $\mathfrak{K}$. If Y is monotone or if u is $\mathfrak{K}_2$, then u is said to be *applicable* (R) to Y.

DEFINITION 4.4. If ϕ is applicable to f and if

$$|\phi:f(t \pm) - \phi:f(t)| = T_{\lambda=0}^{1}[\phi\{(1-\lambda)f(t) + \lambda f(t \pm)\}] \qquad (0 \leqq t \leqq 1),$$

then ϕ is *strictly applicable* to f.

† Throughout the paper we shall consider definitions which involve purely metric properties of a function on $[a, b]$ to one-space to be generalized in the customary manner to functions on $[a, b]$ to n-space.

DEFINITION 4.5. If u is applicable (R) to Y and if for each t on $[0, 1]$ it is true that $u\{t, (1-\lambda)Y(t)+\lambda Y(t\pm)\}$ is monotone in λ for t on $[0, 1]$, then u is said to be *strictly applicable* (R) to Y.

Definitions 4.2 and 4.4 are formulated to facilitate the discussion of the following problem: Suppose $f_n=(X_n+iY_n)-v\rightarrow(X_0+iY_0)=f_0$ and suppose ϕ in CC with $\phi(x+iy)=u(x, y)+iv(x, y)$ for $-\infty<x, y<\infty$. What conditions on ϕ will imply $\phi:f_n-v\rightarrow\phi:f_0$? This is equivalent to asking what condition on ϕ will imply

$$T_{t=0}^1[u\{X_n(t), Y_n(t)\} + iv\{X_n(t), Y_n(t)\}]$$
$$\rightarrow T_{t=0}^1[u\{X_0(t), Y_0(t)\} + iv\{X_0(t), Y_0(t)\}].$$

Definitions 4.3 and 4.5 will be used in connection with convergence in length.

LEMMA 4.3. *Let σ and g be in CR with σ real and non-decreasing and $\sigma(0)=0$, $\sigma(1)=1$. If $f=g:\sigma$ with $f(t\pm)=g\{\sigma(t\pm)\}$ for t on $[0, 1]$, then*

$$\|g\| = \|f\| + T_0^1(\Lambda),$$

where

$$\Lambda(t) = 2^{\operatorname{sgn} t(t-1)}\{T_{\sigma(t-)}^{\sigma(t+)}(g) - |f(t) - f(t-)| - |f(t+) - f(t)|\} \quad (0 \leqq t \leqq 1).$$

Let f_0 and σ_0 be defined on $[-2, 2]$ as follows:

$$\begin{aligned} f_0(t) &= f(t), & \sigma_0(t) &= \sigma(t) & &(0 \leqq t \leqq 1),\\ f_0(t) &= f(1), & \sigma_0(t) &= \sigma(1) & &(1 < t \leqq 2),\\ f_0(t) &= f(0), & \sigma_0(t) &= \sigma(0) & &(-2 \leqq t < 0),\end{aligned}$$

and let

$$\Lambda_0(t) = 2^{-1}\{T_{\sigma_0(t-)}^{\sigma_0(t+)}(g) - |f_0(t) - f_0(t-)| - |f_0(t+) - f_0(t)|\} \quad (-2 \leqq t \leqq 2).$$

The truth of the lemma is clearly equivalent to showing

$$T_0^1(g) = T_{-2}^2(f_0) + T_{-2}^2(\Lambda_0).$$

To do this let $S\equiv(0=s_0<s_1<s_2<\cdots<s_k=1)$ be any partition of $[0, 1]$ and let $H(s', s'')$ be defined for $0\leqq s'\leqq s''\leqq 1$ by

$$H(s', s'') = \sum_{j=1}^{q}|g(p_j) - g(p_{j-1})|,$$

where $p_0=s'$, $p_q=s''$, and q is one more than the number of points of S on $\underset{s}{E}[s'<s<s'']$ and $p_1<p_2<\cdots<p_{q-1}$ are these points (if any). Consider the set

$$E = \underset{t}{E}[H(0, \sigma_0(t)) \leqq T_{-2}^t(f_0) + T_{-2}^t(\Lambda_0)]$$

and denote by E' the set of those points which are left-hand limit points† of E. Noting that $-1\varepsilon E'$ let $P=\sup E'$ and suppose $P<2$ in an attempt to show that $P=2$. Since $P\varepsilon E'$ there is a point P_1 (>-2) of E which is less than P but sufficiently close to P so that $\underset{s}{E}[\sigma_0(P_1)<s<\sigma_0(P-)]$ contains no point of S. It is likewise clear that there exists a point P_2 $(>P$ and $<2)$ such that $\underset{s}{E}[\sigma_0(P+)<s<\sigma_0(P_2)]$ contains no point of S. Now $P_1\varepsilon E$ implies

$$H(0, \sigma_0(P_1)) \leqq T_{-2}^{P_1}(f_0) + T_{-2}^{P_1}(\Lambda_0) \tag{1}$$

and from the fact that Λ_0 vanishes everywhere on $[-2, 2]$ except for a denumerable set it follows that

$$T_{\sigma(P-)}^{\sigma(P+)}(g) - | f_0(P) - f_0(P-) | - | f_0(P+) - f_0(P) | \leqq T_{P_1}^{R}(\Lambda_0) \quad (P < R < P_2)$$

which implies

$$\begin{aligned} H(\sigma_0(P-), \sigma_0(P+)) &\leqq T_{\sigma_0(P-)}^{\sigma_0(P+)}(g) \\ &\leqq | f_0(P) - f_0(P-) | + | f_0(P+) - f_0(P) | + T_{P_1}^{R}(\Lambda_0) \quad (P < R < P_2). \end{aligned} \tag{2}$$

Furthermore the fact that $\underset{s}{E}[\sigma_0(P_1)<s<\sigma_0(P-)]+\underset{s}{E}[\sigma_0(P+)<s<\sigma_0(P_2)]$ contains no point of S combines with the hypothesis of the lemma to yield

$$\begin{aligned} &H(\sigma_0(P_1), \sigma_0(P-)) + H(\sigma_0(P+), \sigma_0(R)) \\ &\qquad = | f_0(P-) - f_0(P_1) | + | f_0(R) - f_0(P+) | \quad (P < R < P_2), \end{aligned} \tag{3}$$

so that upon adding (1), (2), and (3) we obtain, for $P<R<P_2$,

$$\begin{aligned} H(0, \sigma_0(R)) &\leqq H(0, \sigma_0(P_1)) + H(\sigma_0(P_1), \sigma_0(P-)) \\ &\quad + H(\sigma_0(P-), \sigma_0(P+)) + H(\sigma_0(P+), \sigma_0(R)) \\ &\leqq T_{-2}^{P_1}(f_0) + | f_0(P-) - f_0(P_1) | + | f_0(P) - f_0(P-) | \\ &\quad + | f_0(P+) - f_0(P) | + | f_0(R) - f_0(P+) | + T_{-2}^{P_1}(\Lambda_0) + T_{P_1}^{R}(\Lambda_0) \\ &\leqq T_{-2}^{R}(f_0) + T_{-2}^{R}(\Lambda_0). \end{aligned}$$

This establishes P_2 as a point of E' which is a contradiction proving $2=\sup E'$ and hence that

$$\sum_{j=1}^{k} | g(s_j) - g(s_{j-1}) | = H(0, 1) \leqq T_{-2}^{2}(f_0) + T_{-2}^{2}(\Lambda_0),$$

which implies

$$T_0^1(g) \leqq T_{-2}^{2}(f_0) + T_{-2}^{2}(\Lambda_0)$$

since S was arbitrary.

† We define x_0 as a left-hand limit point of E if every interval $[x, x_0]$, where $x<x_0$, contains a point of E as an inner point.

Now let $t_1, t_2, t_3, \cdots$ be a denumerable set of points on $[0, 1]$ which include all points of discontinuity of Λ_0 and define $\Lambda_n(t)=\Lambda_0(t)$ for $t=t_1, t_2, \cdots, t_n$ and $\Lambda_n(t)=0$ for all other t on $[-2, 2]$. Clearly $\Lambda_n(t)\to\Lambda_0(t)$ for all t on $[-2, 2]$ and inasmuch as

$$T_{\sigma_0(t')}^{\sigma_0(t'')}(g) \geqq T_{t'}^{t''}(f_0) \qquad (-2 \leqq t' \leqq t'' \leqq 2),$$

it follows by induction that

$$T_{\sigma_0(t')}^{\sigma_0(t'')}(g) \geqq T_{t'}^{t''}(f_0) + T_{t'}^{t''}(\Lambda_n)$$

for $t' \leqq t''$ with t' and t'' both in the set obtained by deleting the points $t_1, t_2, t_3, \cdots$ from $[-2, 2]$. This of course implies

$$T_0^1(g) \geqq T_{-2}^2(f_0) + T_{-2}^2(\Lambda_n)$$

whence, by semi-continuity

$$T_0^1(g) \geqq T_{-2}^2(f_0) + T_{-2}^2(\Lambda_0)$$

and the proof of the lemma is complete.

LEMMA 4.4. *Let $f=X+iY$ be in BV and denote $T_0^1(f)$ by μ. Let σ be a non-decreasing function in RBV such that $\sigma(0)=0$, $\sigma(1)=1$ and $\mu\sigma(t)=T_0^t(f)$ for t on $[0, 1]$. There exists a function g in BV having the following properties:*

(i) $f = g:\sigma$ *with* $\|f\| = \|g\|$;

(ii) $|g(s_2) - g(s_1)| \leqq \mu|s_2 - s_1|$, $(0 \leqq s_1, s_2 \leqq 1)$, *the sign of equality holding if* $\sigma(t-) \leqq s_1, s_2 \leqq \sigma(t)$ *or* $\sigma(t) \leqq s_1, s_2 \leqq \sigma(t+)$, *where* $0 \leqq t \leqq 1$;

(iii) *if ϕ is any function in CC which is applicable to g, then*

$$\|\phi:g\| = \|\phi:f\| + T_0^1(\Lambda),$$

where

$$\Lambda(t) = 2^{\operatorname{sgn} t(t-1)}\Big\{T_{\lambda=0}^1[\phi\{\lambda f(t-) + (1-\lambda)f(t)\}] - |\phi:f(t) - \phi:f(t-1)| + T_{\lambda=0}^1[\phi\{\lambda f(t) + (1-\lambda)f(t+)\}] - |\phi:f(t+) - \phi:f(t)|\Big\} \quad (0 \leqq t \leqq 1).$$

Define $g=\xi+i\eta$ as follows. Let s_0 be any point on $[0, 1]$ and let $t_0=\inf \underset{t}{E}[\sigma(t)\geqq s_0]$. Now $\sigma(t_0-) \leqq s_0 \leqq \sigma(t_0+)$ and we define

$$g(s_0) = f(t_0) \quad \text{if} \quad s_0 = \sigma(t_0);$$

$$g(s_0) = \frac{f(t_0-)\{\sigma(t_0) - s_0\} + f(t_0)\{s_0 - \sigma(t_0-)\}}{\sigma(t_0) - \sigma(t_0-)} \quad \text{if} \quad \sigma(t_0-) \leqq s_0 < \sigma(t_0);$$

$$g(s_0) = \frac{f(t_0)\{\sigma(t_0+) - s\} + f(t_0+)\{s_0 - \sigma(t_0)\}}{\sigma(t_0+) - \sigma(t_0)} \quad \text{if} \quad \sigma(t_0) < s_0 \leqq \sigma(t_0+).$$

As a consequence of

$$(1) \qquad | f(t'') - f(t') | \leqq \mu | \sigma(t'') - \sigma(t') | \qquad (0 \leqq t', t'' \leqq 1),$$

it is easily verified that $g\{\sigma(t)\} = f(t)$ and $g\{\sigma(t\pm)\} = f(t\pm)$ for t on $[0, 1]$. Combining these last two relations with the definition of g and the relation

$$| f(t \pm) - f(t) | = \mu | \sigma(t \pm) - \sigma(t) | \qquad (0 \leqq t \leqq 1),$$

we obtain

$$(2) \qquad | g(s_2) - g(s_1) | = \mu | s_2 - s_1 |$$

if $\sigma(t-) \leqq s_1, s_2 \leqq \sigma(t)$ or $\sigma(t) \leqq s_1, s_2 \leqq \sigma(t+)$, where $0 \leqq t \leqq 1$. Hence

$$\begin{aligned} T_{\sigma(t-)}^{\sigma(t+)}(g) &= \mu | \sigma(t) - \sigma(t-) | + \mu | \sigma(t+) - \sigma(t) | \\ &= | f(t) - f(t-) | + | f(t+) - f(t) | \qquad (0 \leqq t \leqq 1) \end{aligned}$$

while, on the other hand, from Lemma 3.8 it follows that

$$\begin{aligned} \Lambda(t) = 2^{\operatorname{sgn} t(t-1)} [T_{\sigma(t-)}^{\sigma(t)}(\phi:g) &- | \phi:f(t) - \phi:f(t-) | \\ &+ T_{\sigma(t)}^{\sigma(t+)}(\phi:g) - | \phi:f(t+) - \phi:f(t) |] \\ = 2^{\operatorname{sgn} t(t-1)} [T_{\sigma(t-)}^{\sigma(t+)}(\phi:g) &- | \phi:f(t) - \phi:f(t-) | - | \phi:f(t+) - \phi:f(t) |]. \end{aligned}$$

Viewing the last two relations in the light of Lemma 4.3 establishes $\|g\| = \|f\|$ and $\|\phi:g\| = \|\phi:f\| + T_0^1(\Lambda)$. It also follows that monotonicity of X or Y implies monotonicity of ξ or η respectively so that from Definition 4.2 we conclude ϕ is applicable to g.

To complete the proof of (ii) let $s_1' \leqq s_2'$ be any two numbers on $[0, 1]$ and let $t_1' = \inf \underset{t}{E}[\sigma(t) \geqq s_1']$ and $t_2' = \sup \underset{t}{E}[\sigma(t) \leqq s_2']$. Now $t_1' \leqq t_2'$. If $t_1' = t_2'$ then $\sigma(t_1'-) \leqq s_1' \leqq s_2' \leqq \sigma(t_1'+)$, so that the relation (2) implies the inequality $|g(s_2') - g(s_1')| \leqq \mu |s_2' - s_1'|$. However, if $t_1' < t_2'$ then

$$\sigma(t_1'-) \leqq s_1' \leqq \sigma(t_1'+) \leqq \sigma(t_2'-) \leqq s_2' \leqq \sigma(t_2'+),$$

so that (1) and (2) yield

$$\begin{aligned} | g(s_2') - g(s_1') | &\leqq | g\{\sigma(t_1'+)\} - g(s_1') | + | g\{\sigma(t_2'-)\} - g\{\sigma(t_1'+)\} | \\ &\quad + | g(s_2') - g\{\sigma(t_2'-)\} | \\ &\leqq \mu\{\sigma(t_1'+) - s_1'\} + | f(t_2'-) - f(t_1'+) | + \mu\{s_2' - \sigma(t_2'-)\} \\ &\leqq \mu\{\sigma(t_1'+) - s_1'\} + \mu\{\sigma(t_2'-) - \sigma(t_1'+)\} \\ &\quad + \mu\{s_2' - \sigma(t_2'-)\} = \mu(s_2' - s_1'). \end{aligned}$$

The truth of (ii) is now apparent and the proof of the lemma is complete.

We introduce here the notion of pseudo-absolute continuity.

DEFINITION 4.6. A function f defined on $[a, b]$ is said to be *pseudo-absolutely continuous* there if corresponding to every $\epsilon(>0)$ there exists a $\delta(>0)$ and a finite point set E such that if $\{[a_n, b_n]\}$ is any denumerable set of non-overlapping intervals on $[a, b]$ with $E\cdot\sum_{n=1}^{\infty}[a_n, b_n]$ empty and $\sum_{n=1}^{\infty}|b_n-a_n|<\delta$, then $\sum_{n=1}^{\infty}|f(b_n)-f(a_n)|<\epsilon$.

We observe that a pseudo-a.c. function is of b.v. and is expressible as the sum of an a.c. function and a singular function of the saltus type (see Definition 6.2 below).

LEMMA 4.5. *If f is a pseudo-absolutely continuous point in BV and if ϕ is any function in CC which is applicable to f, then $\phi{:}g$ is a pseudo-absolutely continuous point in BV.*

There is no loss in generality in assuming ϕ to be real valued.

Let $X+iY=f$. From Definitions 4.1 and 4.2 there exist functions A and B on one-space to n_1- and n_2-spaces respectively which are a.c. on every finite interval and a function U on n_3-space $(n_3=n_1+n_2)$ to one-space, all of whose first partial derivatives are continuous, such that

$$\phi(x + iy) = U\{A(x) \circ B(y)\} \qquad (-\infty < x, y < \infty)$$

with A satisfying a Lipschitz condition on every finite interval if X is not monotone and B satisfying a Lipschitz condition on every finite interval if Y is not monotone. Let

$$C(t) = A\{X(t)\} \circ B\{Y(t)\} \qquad (0 \leqq t \leqq 1).$$

Readily seen is the pseudo-absolute continuity† of C. Let S be a sphere in n_3-space which includes the range of C. Now U satisfies a Lipschitz condition on S; i.e., there exists a constant $M(>0)$ such that if γ and γ' are any two points of S, then

$$|U(\gamma') - U(\gamma)| \leqq M\{\text{Euclidean distance between } \gamma \text{ and } \gamma'\},$$

so that, since $\phi{:}f(t)=U\{C(t)\}$ for t on $[0, 1]$, the pseudo-absolute continuity of $\phi{:}f$ becomes apparent.

Since a continuous pseudo-absolutely continuous function is a.c. we have

COROLLARY 4.1. *If f is an a.c. point in BV and if ϕ is any function in CC which is applicable to f, then $\phi{:}g$ is an a.c. point in BV.*

LEMMA 4.6. *If $\{g_n\}$ is a sequence of points in BV satisfying the relation*

$$|g_n(t_2) - g_n(t_1)| \leqq M|t_2 - t_1| \qquad (0 \leqq t_1, t_2 \leqq 1; n = 0, 1, 2, \cdots).$$

† See footnote on page 59.

and if ϕ *is any function in CC which is applicable to* g_n *for* $n=0, 1, 2, \cdots$, *then the relation* $\|g_n-g_0\|\to 0$ *implies* $\|\phi:g_n-\phi:g_0\|\to 0$.

There is no loss in generality in assuming ϕ real valued. Let $\{g_{n,1}\}$ be any subsequence of $\{g_n\}$ wherein $g_{0,1}=g_0$. To prove $\|\phi:g_n-\phi:g_0\|\to 0$ it is merely necessary to establish the existence of a sequence $\{g_{n,2}\}$ which is a subsequence of $\{g_{n,1}\}$, such that $\|\phi:g_{n,2}-\phi:g_0\|\to 0$. First we note that of the sequence $\{g_{n,1}\}$ there exists a subsequence $\{g_{n,2}\}$ wherein $g_{0,2}=g_0$ and $X_n+iY_n=g_{n,2}$ for $n=0, 1, 2, \cdots$, which enjoys one of the following four properties:

(i) X_1 and Y_1 are both not monotone;
(ii) X_n is monotone ($n=0, 1, 2, \cdots$) and Y_1 is not;
(iii) X_n and Y_n are monotone ($n=0, 1, 2, \cdots$);
(iv) X_1 is not monotone and Y_n is monotone ($n=0, 1, 2, \cdots$).

Now since ϕ is applicable to $\{g_{1,2}\}$, there exist functions A and B on one-space to n_1- and n_2-spaces respectively and a real function U defined on n_3-space ($n_3=n_1+n_2$), all of whose first partial derivatives are continuous, such that

$$\phi(x+iy) = U\{A(x)\circ B(y)\} \qquad (-\infty < x, y < \infty)$$

and such that C_n defined by $C_n(t)=A\{X_n(t)\}\circ B\{Y_n(t)\}$ $(0\leq t\leq 1)$, is an a.c. function on $[0, 1]$ to n_3-space not only for $n=1$, but for $n=0, 1, 2, \cdots$. Thus upon defining

$$(C_{n,1}(t), C_{n,2}(t), \cdots, C_{n,n_3}(t)) = C_n(t) \text{ and } H_n(t) = U\{C_n(t)\}$$
$$(0 \leq t \leq 1, n = 0, 1, 2, \cdots),$$

we conclude H_n is an a.c. point in RBV with $H_n=\phi:g_{n,2}$ for $n=0, 1, 2, \cdots$. Also follows the existence of continuous functions $D_1, D_2, \cdots, D_{n_3}$ on n_3-space to one-space such that (for $n=0, 1, 2, \cdots$)

$$H_n'(t) = \sum_{p=1}^{n_3} D_p\{C_n(t)\}C_{n,p}'(t) \qquad \text{(almost all } t \text{ on } [0, 1]).$$

Since g_0 is continuous, we conclude $(X_n+iY_n)-uv\to(X_0+iY_0)$ and hence $C_n(t)\to C_0(t)$ uniformly for t on $[0, 1]$, so that

$$\lim_{n\to\infty} D_p\{C_n(t)\} = D_p\{C_0(t)\} \qquad (p = 1, 2, \cdots, n_3)$$

uniformly for t on $[0, 1]$. Combining this with the relation

$$\lim_{n\to\infty} \int_0^1 | C_{n,p}'(t) - C_{0,p}'(t) | \, dt = 0 \quad (p = 1, 2, \cdots, n_3),$$

which is a corollary of Lemma 4.1 and Lemma 4.2, yields

$$T_0^1(H_n - H_0) = \int_0^1 | H_n'(t) - H_0'(t) | \, dt \to 0.$$

Hence, since $H_n(0) \to H_0(0)$, we conclude $\|\phi : g_{n,2} - \phi : g_0\| \to 0$. This completes the proof.

With this background we now turn to the proofs of the following two theorems.

THEOREM 4.1. *If ϕ is applicable to f_n for $n = 0, 1, 2, \cdots$ and in addition if ϕ is strictly applicable to f_0, then the relation $f_n - v \to f_0$ implies*

$$\phi : f_n - v \to \phi : f_0.$$

THEOREM 4.2. *If ϕ is applicable to f_n for $n = 0, 1, 2, \cdots$, then the relation $f_n - uv \to f_0$ implies*

$$\phi : f_n - uv \to \phi : f_0.$$

Let $\mu_n = T_0^1(f_n)$, $(n = 0, 1, 2, \cdots)$. From Lemma 4.4 we conclude the existence (for $n = 0, 1, 2, \cdots$) of functions g_n and σ_n in CR having the following properties:

(i) σ_n is a non-decreasing function with $\sigma_n(0) = 0$, $\sigma_n(1) = 1$ and $\mu_n \sigma_n(t) = T_0^t(f_n)$ for $0 \leqq t \leqq 1$;

(ii) $f_n = g_n : \sigma_n$ with $\|f_n\| = \|g_n\|$;

(iii) $|g_n(s_2) - g_n(s_1)| \leqq \mu_n |s_2 - s_1|$, $(0 \leqq s_1, s_2 \leqq 1)$, the sign of equality holding if $\sigma_n(t-) \leqq s_1, s_2 \leqq \sigma_n(t)$ or $\sigma_n(t) \leqq s_1, s_2 \leqq \sigma_n(t+)$, where $0 \leqq t \leqq 1$;

(iv) ϕ is applicable to g_n and

$$\|\phi : g_n\| = \|\phi : f_n\| + T_0^1(\Lambda_n),$$

where

$$\Lambda_n(t) = 2^{\operatorname{sgn} t(t-1)} \Big\{ T_{\lambda=0}^1 [\phi\{\lambda f_n(t-) + (1-\lambda) f_n(t)\}] - |\phi : f_n(t) - \phi : f_n(t-)| \\ + T_{\lambda=0}^1 [\phi\{\lambda f_n(t) + (1-\lambda) f_n(t+)\}] - |\phi : f_n(t+) - \phi : f_n(t)| \Big\} \\ (0 \leqq t \leqq 1).$$

We divide the remainder of the proof of Theorems 4.1 and 4.2 into three parts.

PART I. *$\phi : g_n$ is a.c. for each $n = 0, 1, 2, \cdots$ and $\|\phi : g_n - \phi : g_0\| \to 0$.*

From (i) and Lemma 3.1 follows the relation, $\mu_n \sigma_n(t) \to \mu_0 \sigma_0(t)$ for t on $[0, 1]$ so that

$$| g_n(\sigma_n(t)) - g_n(\sigma_0(t)) | \leqq | \mu_n \sigma_n(t) - \mu_n \sigma_0(t) | \to 0$$

since $\mu_n \to \mu_0$. From (ii) follows the relation $g_n\{\sigma_n(t)\} \to g_0\{\sigma_0(t)\}$ for t on $[0, 1]$, so that upon combining the above relations we conclude

$$g_n\{\sigma_0(t)\} \to g_0\{\sigma_0(t)\} \qquad (0 \leqq t \leqq 1).$$

Now let s_0 be any point on $[0, 1]$ and note that there exists a point t_0 on $[0, 1]$ for which $\sigma_0(t_0-) \leqq s_0 \leqq \sigma_0(t_0+)$, so that either $s_0 = \sigma_0(t_0)$, or $\sigma_0(t_0-) \leqq s_0 < \sigma_0(t_0)$, or $\sigma_0(t_0) < s_0 \leqq \sigma_0(t_0+)$. If $s_0 = \sigma_0(t_0)$ then follows immediately the conclusion $g_n(s_0) \to g_0(s_0)$. Supposing now that $\sigma_0(t_0-) \leqq s_0 < \sigma_0(t_0)$, let $0 < t_1 < t_2 < t_3 < \cdots$ with $t_p \to t_0$ and denote by P any limit point of the sequence $\{g_n(s_0)\}$. Also let $s_0' = \sigma_0(t_0-)$ and $s_0'' = \sigma_0(t_0)$, $P_0' = g_0(s_0')$, $P_0 = g_0(s_0)$, and $P_0'' = g_0(s_0'')$. Thus

$$| g_n(s_0) - g_n\{\sigma_0(t_p)\} | \leqq \mu_n\{s_0 - \sigma_0(t_p)\},$$
$$| g_n(s_0'') - g_n(s_0) | \leqq \mu_n\{s_0'' - s_0\} \qquad (n = 0, 1, 2, \cdots; p = 1, 2, 3, \cdots).$$

Letting $n \to \infty$ and then $p \to \infty$ and then using (iii) establishes

$$| P - P_0' | \leqq \mu_0\{s_0 - s_0'\} = | P_0 - P_0' |$$
$$| P_0'' - P | \leqq \mu_0\{s_0'' - s_0\} = | P_0'' - P_0 |.$$

Adding and using (iii) again yields

$$| P_0'' - P_0' | \leqq | P - P_0' | + | P_0'' - P | \leqq \mu_0(s_0'' - s_0') = | P_0'' - P_0' |$$

which implies equality in the last three relations which in turn implies $P_0 = P$. Thus $g_n(s_0) \to g_0(s_0)$ if $\sigma_0(t_0-) \leqq s_0 < \sigma_0(t_0)$. A similar proof of this relation holds if $\sigma_0(t_0) < s_0 \leqq \sigma_0(t_0+)$, so that finally it is established that $g_n(s) \to g_0(s)$ for s on $[0, 1]$.

From (iii) we conclude g_n is a.c. with $|g_n'(s)| \leqq \mu_n$, for $0 \leqq s \leqq 1$; $n = 0, 1, 2, \cdots$. Hence (for $n = 0, 1, 2, \cdots$)

$$\int_0^1 \Big| \mu_n - | g_n'(s) | \Big| ds = \int_0^1 \{\mu_n - | g_n'(s) | \} ds = 0,$$

so that $\mu_n = |g_n'(s)|$ for almost all s on $[0, 1]$. Applying Theorem 3.3 yields $\|g_n - g_0\| \to 0$. Letting $M (>0)$ be such that $\mu_n < M$ for $n = 0, 1, 2, \cdots$ we see from (iii) that

$$| g_n(s_2) - g_n(s_1) | \leqq M | s_2 - s_1 | \qquad (0 \leqq s_1, s_2 \leqq 1; n = 0, 1, 2, \cdots),$$

so that Lemma 4.6 and Corollary 4.1 complete the proof of Part I.

Part II. (Proof of Theorem 4.1.) Since ϕ is strictly applicable to f_0 it is apparent from (iv) that $\Lambda_0(t) = 0$, $(0 \leqq t \leqq 1)$, and hence

$$\|\phi : g_n\| \geqq \|\phi : f_n\| \qquad (n = 0, 1, 2, \cdots),$$

the equality holding if $n=0$. Thus $\phi:f_n$ is in BV for $n=0, 1, 2, \cdots$ and further, since $\phi:f_n(t)\to\phi:f_0(t)$ for t on $[0, 1]$, we have

$$0 = \limsup_{n\to\infty} \|\phi:g_n - \phi:g_0\| \geqq \limsup_{n\to\infty} \|\phi:g_n\| - \|\phi:g_0\|$$
$$\geqq \limsup_{n\to\infty} \|\phi:f_n\| - \|\phi:f_0\| \geqq \liminf_{n\to\infty} \|\phi:f_n\| - \|\phi:f_0\| \geqq 0$$

by the semi-continuity property of total variation.

The proof of Part II is now complete.

PART III. (Proof of Theorem 4.2.) Since $\|\phi:g_n\| \geqq \|\phi:f_n\|$ we conclude as before that $\phi:f_n$ is a point of BV for each $n=0, 1, 2, \cdots$. Since, however, $f_n - uv\to f_0$ it becomes clear that in this case

$$f_n(t \pm) \to f_0(t \pm) \quad \text{uniformly for } t \text{ on } [0, 1],$$

which used in connection with Lemma 4.6 readily establishes

$$\Lambda_n(t) \to \Lambda_0(t) \qquad (0 \leqq t \leqq 1),$$

so that $\liminf_{n\to\infty} T_0^1(\Lambda_n) - T_0^1(\Lambda_0) \geqq 0$. Hence

$$0 = \limsup_{n\to\infty} \|\phi:g_n - \phi:g_0\| \geqq \limsup_{n\to\infty} \|\phi:g_n\| - \|\phi:g_0\|$$
$$\geqq \limsup_{n\to\infty} \|\phi:f_n\| - \|\phi:f_0\| + \liminf_{n\to\infty} T_0^1(\Lambda_n) - T_0^1(\Lambda_0)$$
$$\geqq \limsup_{n\to\infty} \|\phi:f_n\| - \|\phi:f_0\| \geqq \liminf_{n\to\infty} \|\phi:f_n\| - \|\phi:f_0\| \geqq 0.$$

Since $\phi:f_n(t)\to\phi:f_0(t)$ uniformly for t on $[0, 1]$, the proof of Part III is now complete.

As a corollary we have

COROLLARY 4.2. *If the function ϕ is in CC with $\phi(x+iy)=u(x, y)+iv(x, y)$ for $-\infty<x, y<\infty$, where u, v have continuous first partial derivatives, then the relation $f_n-v\to f_0$ with f_0 continuous implies*

$$\phi:f_n - v \to \phi:f_0.$$

If the transformation is that of raising to a positive integer power, Theorems 4.1 and 4.2 lead to

COROLLARY 4.3. *Let k be a positive integer. If $f_n-v\to f_0$ and if corresponding to each t on $[0, 1]$ there is a ray through the origin (of the complex plane) on which lie the points $f_0(t)$, $f_0(t+)$, and $f_0(t-)$, then*

$$f_n^k - v \to f_0^k.$$

COROLLARY 4.4. *If $f_n-uv\to f_0$, then $f_n^k-uv\to f_0^k$.*

Before concluding this section the following remark seems in order. If ϕ is in CC and satisfies a Lipschitz condition on every bounded set in the complex plane, then $f\epsilon BV$ implies $\phi:f\epsilon BV$. Hence it is natural to inquire into the truth of the following statement.

If $\phi\epsilon CC$ and satisfies a Lipschitz condition on every finite set and if f_0 is continuous, then the relation $f_n-v\to f_0$ implies $\phi:f_n-v\to\phi:f_0$.

That the statement is not true is illustrated by the following example. Let $f_n(t)=t+i/n$ for $0\leqq t\leqq 1$; $n=1, 2, 3, \cdots$. Let $f_0(t)=t$. It is clear that $f_n-v\to f_0$. Define

$$u(x, y) = \frac{\sin n^4\pi x}{n^4\pi} + \left\{\frac{\sin (n-1)^4\pi x}{(n-1)^4\pi} - \frac{\sin n^4\pi x}{n^4\pi}\right\} \sin^2\left\{\frac{\pi n(n-1)}{2}\left(y-\frac{1}{n}\right)\right\},$$

if $1/n\leqq y<1/(n-1)$, where $n\geqq 2$;

$$u(x, y) = \frac{\sin \pi x}{\pi} \quad \text{if} \quad y \geqq 1; \qquad u(x, y) = 0 \quad \text{if} \quad y \leqq 0;$$

and let $\phi(x+iy)=u(x, y)$ for $-\infty<x, y<\infty$. Since the first partial derivatives of u exist everywhere and are dominated by 3, it follows that ϕ satisfies a Lipschitz condition on the complex plane. But

$$\|\phi:f_n\| = \int_0^1 |\cos n^4\pi x|\, dx \geqq \int_0^1 \cos^2 n^4\pi x\, dx = \tfrac{1}{2} \qquad (n = 1, 2, 3, \cdots)$$

with $\|\phi:f_0\|=0$ so that it is not true that $\phi:f_n-v\to\phi:f_0$. As a rough appraisal of the generality of Theorems 4.1 and 4.2 it is interesting to note that a function in CC may be applicable to f_n for $n=0, 1, 2,$ without satisfying a Lipschitz condition on every bounded set in the complex plane.

5. **Convergence in length.** As our first application of preceding results we have the following theorem, which is a result obtained in a different way in AL.

THEOREM 5.1. *The relation $Y_n-l\to Y_0$ implies $Y_n-v\to Y_0$.*

Let $\phi(x+iy)=y$, $(-\infty<x, y<\infty)$. Now ϕ is strictly applicable to $I+iY_n$ for $n=0, 1, 2, \cdots$ and since $(I+iY_n)-v\to(I+iY_0)$ we conclude that $\phi:(I+iY_n)-v\to\phi:(I+iY_0)$. Hence the theorem is established.

COROLLARY 5.1. *The relation $Y_n-ul\to Y_0$ implies $Y_n-uv\to Y_0$.*

We introduce here the notion of a singular function.

DEFINITION 5.1. If f is of b.v. on $[a, b]$ with $f'(t)=0$ for almost all t on $[a, b]$ then f is said to be *singular* on $[a, b]$.

A well known property of singular functions is this: Let f be a singular function in BV and assume g an a.c. function in BV. Then

$$T_0^1(f+g) = T_0^1(f) + T_0^1(g),$$

so that if either $f(0)=0$ or $g(0)=0$, it is clear that $\|f+g\| = \|f\| + \|g\|$. We are now prepared to prove the following

THEOREM 5.2. *If Y_0 is a singular function in RBV, then the relations $Y_n - l \rightarrow Y_0$ and $Y_n - v \rightarrow Y_0$ are equivalent.*

Supposing $Y_n - v \rightarrow Y_0$ we deduce the relation

$$\|I + iY_0\| \leqq \liminf_{n\to\infty} \|I + iY_n\| \leqq \limsup_{n\to\infty} \|I + iY_n\| \leqq \limsup_{n\to\infty} \|Y_n\| + \|I\|$$
$$= \|I\| + \|Y_0\| = \|I\| + \|iY_0\| = \|I + iY_0\|,$$

which proves $Y_n - l \rightarrow Y_0$.

Application of Theorem 5.1 completes the proof.

THEOREM 5.3. *If u is applicable (R) to Y_n for $n=0, 1, 2, \cdots$ and in addition if u is strictly applicable (R) to Y_0, then the relation $Y_n - l \rightarrow Y_0$ implies*

$$(u|Y_n) - l \rightarrow (u|Y_0).$$

Define $\phi(x+iy) = x + iu(x, y)$ for $-\infty < x, y < \infty$. Now ϕ is applicable to $(I+iY_n)$ for $n=0, 1, 2, \cdots$. It is also strictly applicable to $(I+iY_0)$. Hence $\phi:(I+iY_n) - v \rightarrow \phi:(I+iY_0)$ or $\{I+i(u|Y_n)\} - v \rightarrow \{I+i(u|Y_0)\}$.

THEOREM 5.4. *If X_0 is a.c., then the relations $X_n - l \rightarrow X_0$ and $Y_n - l \rightarrow Y_0$ imply the relations*†

$$(X_n + Y_n) - l \rightarrow (X_0 + Y_0),$$
$$X_n Y_n - l \rightarrow X_0 Y_0.$$

Let Ψ be a.c. on $\underset{x}{E}[-\infty < x < \infty]$ with $\Psi(x) = X_0(x)$ for x on $[0, 1]$. Define $u_1(x, y) = y - \Psi(x)$, $u_2(x, y) = y + \Psi(x)$; $-\infty < x, y < \infty$. Thus by Theorem 5.3 we have $(u_1|X_n) - l \rightarrow \theta$, which implies (Theorem 5.2) that $\|X_n - X_0\| \rightarrow 0$ (see §2 for definition of θ). Hence

$$\|I + i(Y_n + X_n - X_0)\| \leqq \|I + iY_n\| + \|X_n - X_0\| \rightarrow \|I + iY_0\|,$$

whence, by using the semi-continuity property of total variation, we deduce

$$(Y_n + X_n - X_0) - l \rightarrow Y_0.$$

---

† This is a generalization of Theorem 6 in AL.
If the assumption that X_0 is a.c. is deleted the theorem ceases to be true. See AL, page 23.

This gives, in view of Theorem 5.3,

$$[u_2|(Y_n + X_n - X_0)] - l \rightarrow (u_2|Y_0) \quad \text{or} \quad (Y_n + X_n) - l \rightarrow (Y_0 + X_0).$$

Now by Lemma 3.2 we have

$$\|Y_n \cdot (X_n - X_0)\| \leqq \|Y_n\| \cdot \|X_n - X_0\| \rightarrow 0,$$

so that upon defining $u_3(x, y) = y\Psi(x)$, $-\infty < x, y < \infty$, it is seen by Theorem 5.3 that

$$(u_3|Y_n) - l \rightarrow (u_3|Y_0) \quad \text{or} \quad X_0 Y_n - l \rightarrow X_0 Y_0$$

and since (Theorem 5.2)

$$(X_n Y_n - X_0 Y_n) = [Y_n \cdot (X_n - X_0)] - l \rightarrow \theta,$$

we conclude upon adding, θ being a.c., that $X_n Y_n - l \rightarrow X_0 Y_0$.

Lemma 5.1. *Let X be an a.c. point in the space RBV. The relation $(cI + Y_n) - v \rightarrow (cI + Y_0)$ for all real c implies $(X + Y_n) - v \rightarrow (X + Y_0)$.*

Let $[a, b]$ be a subinterval of $[0, 1]$. From Lemma 3.1 it follows that

$$T_{t=a}^{b}(c_1 + c_2 t + Y_n(t)) \rightarrow T_{t=a}^{b}(c_1 + c_2 t + Y_0(t))$$

for all real c_1 and c_2, whence we conclude

$$(\beta + Y_n) - v \rightarrow (\beta + Y_0),$$

where β is any polygonal function in RBV.

Let $\{\beta_p\}$ be a sequence of polygonal functions in RBV such that as $p \rightarrow \infty$ $\beta_p - l \rightarrow X$. From Theorems 5.4 and 5.1 follow the relations $\|\beta_p - X\| \rightarrow 0$ and $(\beta_p + Y_0) - v \rightarrow (X + Y_0)$ as $p \rightarrow \infty$. Hence

$$\begin{aligned} \|X + Y_0\| &\leqq \liminf_{n\rightarrow\infty} \|X + Y_n\| \leqq \limsup_{n\rightarrow\infty} \|X + Y_n\| \\ &\leqq \limsup_{n\rightarrow\infty} \|\beta_p + Y_n\| + \lim_{n\rightarrow\infty} \|X - \beta_p\| \\ &= \|\beta_p + Y_0\| + \|X - \beta_p\| \rightarrow \|X + Y_0\| \quad \text{as} \quad p \rightarrow \infty \end{aligned}$$

and the lemma is proved.

Theorem 5.5. *The relation† $(cI + Y_n) - v \rightarrow (cI + Y_0)$ for all real numbers c and the relation $Y_n - l \rightarrow Y_0$ are equivalent.*

Let $\alpha(t) = \int_0^t Y_0'(s)ds$ and $\beta(t) = Y_0(t) - \alpha(t)$ for t on $[0, 1]$. From the preceding lemma follows

† We are indebted to Professor E. J. McShane for raising the question as to whether the relation $(cI + Y_n) - v \rightarrow (cI + Y_0)$ for all real numbers c implies $Y_n - l \rightarrow Y_0$.

$$(Y_n - \alpha) - v \rightarrow \beta.$$

Since β is singular we deduce from Theorems 5.2 and 5.4 that

$$(Y_n - \alpha) - l \rightarrow \beta \quad \text{and} \quad Y_n - l \rightarrow (\alpha + \beta) = Y_0.$$

From Theorem 5.4 the converse follows immediately.

If Y_0 is a.c. then it appears, as a consequence of Theorems 5.1 and 5.4, that the relation $Y_n - l \rightarrow Y_0$ implies

$$\int_0^1 | Y_n'(t) - Y_0'(t) | \, dt \rightarrow 0.$$

However, if Y_0 is not a.c., this conclusion need not be true. It is true, nevertheless, that Y_n' converges to Y_0' in a manner intermediate between convergence in the mean and convergence in measure. To characterize this type of convergence we introduce the following definitions which may have some intrinsic interest.

Definition 5.2. If f_n is measurable on a set E for $n=1, 2, 3, \cdots$ and if corresponding to every $\epsilon > 0$ there exists a measurable set $E_1 \subset E$ of measure $> |E| - \epsilon$ such that

$$\lim_{m,n\rightarrow\infty} \int_{E_1} | f_m(t) - f_n(t) | \, dt = 0,$$

then $\{f_n\}$ is said to be convergent *almost in the mean* on E.

Definition 5.3. If f_n is measurable on a set E for $n=0, 1, 2, \cdots$ and if corresponding to every $\epsilon > 0$ there exists a measurable set $E_1 \subset E$ of measure $> |E| - \epsilon$ such that

$$\int_{E_1} | f_n(t) - f_0(t) | \, dt \rightarrow 0,$$

then f_n is said to converge *almost in the mean to* f_0 on E.

If $\{f_n\}$ is convergent almost in the mean on E then it is easily seen that there exists a function f_0 defined on E such that f_n converges almost in the mean to f_0 on E.

Definition 5.4. By $f_n - \mu \rightarrow f_0$ is meant this: f_n is in CR for $n=0, 1, 2, \cdots$ and f_n converges almost in the mean to f_0 on $[0, 1]$.

Theorem 5.6. *The relation* $Y_n - l \rightarrow Y_0$ *implies and is implied by the two relations* $Y_n - v \rightarrow Y_0$ *and* $Y_n' - \mu \rightarrow Y_0'$.

We have already seen (Theorem 5.1) that $Y_n - l \rightarrow Y_0$ implies $Y_n - v \rightarrow Y_0$. We now propose to show that $Y_n - l \rightarrow Y_0$ implies $Y_n' - \mu \rightarrow Y_0'$.

Define

$$\alpha(t) = \int_0^t Y_0(s)ds, \qquad \beta(t) = Y_0(t) - \alpha(t) \text{ for } t \text{ on } [0, 1],$$

and let ϵ be any positive number. The singularity of β implies, as is well known, the existence for each $m=1, 2, 3, \cdots$ of non-overlapping intervals $[a_{m,1}, b_{m,1}]$, $[a_{m,2}, b_{m,2}]$, $\cdots$, $[a_{m,N_m}, b_{m,N_m}]$ contained in $[0, 1]$ such that

$$\sum_{j=1}^{N_m} T_{a_{m,j}}^{b_{m,j}}(\beta) < \frac{1}{m}, \qquad |A_m| > 1 - \frac{\epsilon}{2^m},$$

where $A_m = \sum_{j=1}^{N_m} [a_{m,j}, b_{m,j}]$. The absolute continuity of the function α implies that $(Y_n - \alpha) - l \rightarrow \beta$ which in turn implies $(Y_n - \alpha) - v \rightarrow \beta$. Letting $A = A_1 A_2 A_3 \cdots$ we conclude from Lemmas 3.1 and 3.7 that

$$\limsup_{n\to\infty} \int_A | Y_n'(t) - Y_0'(t) | dt = \limsup_{n\to\infty} \int_A | Y_n'(t) - \alpha'(t) | dt$$

$$\leqq \limsup_{n\to\infty} \int_{A_m} | Y_n'(t) - \alpha'(t) | dt = \limsup_{n\to\infty} \sum_{j=1}^{N_m} \int_{a_{m,j}}^{b_{m,j}} | Y_n'(t) - \alpha'(t) | dt$$

$$\leqq \limsup_{n\to\infty} \sum_{j=1}^{N_m} T_{a_{m,j}}^{b_{m,j}}(Y_n - \alpha) = \sum_{j=1}^{N_m} T_{a_{m,j}}^{b_{m,j}}(\beta) < \frac{1}{m} \qquad (m = 1, 2, 3, \cdots)$$

which implies $\int_A | Y_n'(t) - Y_0'(t) | dt \rightarrow 0$. Clearly $|A| > 1 - \epsilon$ so that from Definition 5.4 follows the relation $Y_n' - \mu \rightarrow Y_0'$.

Let us assume now that $Y_n - v \rightarrow Y_0$ and $Y_n' - \mu \rightarrow Y_0'$. Define

$$\alpha_n(t) = \int_0^t Y_n(s)ds, \quad \beta_n(t) = Y_n(t) - \alpha_n(t) \quad \text{for } 0 \leqq t \leqq 1; n = 0, 1, 2, \cdots,$$

and let $\epsilon > 0$. There exists a set $E \subset [0, 1]$ such that $|E| > 1 - \epsilon$ and

$$\int_E | \alpha_n'(t) - \alpha_0'(t) | dt = \int_E | Y_n'(t) - Y_0'(t) | dt \rightarrow 0.$$

Denoting by E' the complement of E with respect to $[0, 1]$ it is seen that the last relation combines with

$$\int_E | \alpha_n'(t) | dt + \int_{E'} | \alpha_n'(t) | dt + T_0^1(\beta_n) - \int_E | \alpha_0'(t) | dt$$

$$- \int_{E'} | \alpha_0'(t) | dt - T_0^1(\beta_0) = \{ T_0^1(Y_n) - T_0^1(Y_0) \} \rightarrow 0$$

to yield

$$\left\{ \int_{E'} | \alpha_n'(t) | dt + T_0^1(\beta_n) - \int_{E'} | \alpha_0'(t) | dt - T_0^1(\beta_0) \right\} \rightarrow 0.$$

Thus from the relation

$$
\begin{aligned}
T^1_{t=0}\{t + iY_n(t)\} &- T^1_{t=0}\{t + iY_0(t)\} \\
&= T^1_{t=0}\{t + i\alpha_n(t)\} + T^1_0(\beta_n) - T^1_{t=0}\{t + i\alpha_0(t)\} - T^1_0(\beta_0) \\
&= \int_E | 1 + i\alpha_n'(t) | \, dt + \int_{E'} | 1 + \alpha_n'(t) | \, dt + T^1_0(\beta_n) \\
&\quad - \int_E | 1 + i\alpha_0'(t) | \, dt - \int_{E'} | 1 + \alpha_0'(t) | \, dt - T^1_0(\beta_0) \\
&\leqq \int_E | \alpha_n'(t) - \alpha_0'(t) | \, dt + \int_{E'} | \alpha_n'(t) | \, dt + T^1_0(\beta_n) \\
&\quad - \int_{E'} | \alpha_0'(t) | \, dt - T^1_0(\beta_0) + 2 | E' |,
\end{aligned}
$$

which holds for $n=1, 2, 3, \cdots$ we conclude

$$
\begin{aligned}
0 &\leqq \liminf_{n\to\infty} T^1_{t=0}\{t + iY_n(t)\} - T^1_{t=0}\{t + iY_0(t)\} \\
&\leqq \limsup_{n\to\infty} T^1_{t=0}\{t + iY_n(t)\} - T^1_{t=0}\{t + iY_0(t)\} \leqq 2 | E' | < 2\epsilon.
\end{aligned}
$$

The arbitrariness of ϵ completes the proof.

COROLLARY 5.2. *The relation $Y_n - l \to Y_0$ implies and is implied by the two relations $Y_n - v \to Y_0$ and Y_n' converges in measure to Y_0' on* $[0, 1]$.

Convergence in measure implies almost convergence in the mean of a subsequence.

An immediate consequence is

THEOREM 5.7. *Let $X_n - l \to X_0$ and $Y_n - l \to Y_0$. Then a necessary and sufficient condition for $(X_n + Y_n) - l \to (X_0 + Y_0)$ is that $(X_n + Y_n) - v \to (X_0 + Y_0)$; furthermore, a necessary and sufficient condition for $X_nY_n - l \to X_0Y_0$ is that $X_nY_n - v \to X_0Y_0$.*

THEOREM 5.8. *Let Y_n be in RBV with $P_n(t)$ and $N_n(t)$ denoting the positive and negative variations of Y_n on* $[0, 1]$, $(n=0, 1, 2, \cdots; 0 \leqq t \leqq 1)$. *Then the relation $Y_n - l \to Y_0$ implies the relations $P_n - l \to P_0$ and $N_n - l \to N_0$.*

To prove this theorem verify first the relations† $P_n - v \to P_0$, $N_n - v \to N_0$, and

$$
\tfrac{1}{2}(P_n - N_n) - l \to \tfrac{1}{2}(P_0 - N_0), \qquad \tfrac{1}{2}(P_n + N_n) - l \to \tfrac{1}{2}(P_0 + N_0).
$$

The desired conclusions are now immediate consequences of Theorem 5.7.

† AC, Theorem 1; AL, Corollary to Theorem 1.

6. **Uniform convergence in length.** Theorem 4.2 combined with the methods used in proving Theorem 5.3 yields

THEOREM 6.1. *If u is applicable (R) to Y_n for $n=0, 1, 2, \cdots$, then the relation $Y_n - ul \rightarrow Y_0$ implies*

$$(u|Y_n) - ul \rightarrow (u|Y_0).$$

We now recall the concept of an elementary step-function and of a singular function of the saltus type.

DEFINITION 6.1. A function f defined on $[a, b]$ is said to be an *elementary step-function* there if there exists a real number c on $[a, b]$ and complex numbers $\gamma_1, \gamma_2, \gamma_3$, such that

$$f(t) = \gamma_1, \text{ if } 0 \leqq t < c; \; f(c) = \gamma_2; \; f(t) = \gamma_3 \text{ if } c < t \leqq b.$$

DEFINITION 6.2. A function f of b.v. on $[a, b]$ is said to be a *singular function of the saltus type* on $[a, b]$ if there exist elementary step-functions $f_1, f_2, f_3, \cdots$ defined on $[a, b]$ such that

$$f(t) = \sum_{n=1}^{\infty} f_n(t), \quad a \leqq t \leqq b; \qquad T_a^b(f) = \sum_{n=1}^{\infty} T_a^b(f_n).$$

It is readily seen that a singular function of the saltus type is singular, though we shall not make explicit use of this property. From the definition of pseudo-absolute continuity it follows that if f is a pseudo-absolutely continuous function in BV, then there exists an a.c. function α in BV and a singular function β of the saltus type in BV such that $f = \alpha + \beta$.

LEMMA 6.1. *If β is an elementary step-function in BV, then the relation $f_n - uv \rightarrow f_0$ implies the relation*

$$(f_n + \beta) - uv \rightarrow (f_0 + \beta).$$

There exist a real number c on $[0, 1]$ and complex numbers $\gamma_1, \gamma_2, \gamma_3$, such that $\beta(t) = \gamma_1$ if $0 \leqq t < c$, $\beta(c) = \gamma_2$, and $\beta(t) = \gamma_3$ if $c < t \leqq 1$, so that as a consequence of Lemma 3.1 it may be deduced that

$$\begin{aligned}
T_0^{c-}(f_n + \beta) &= T_{t=0}^{c-}\{f_n(t) + \gamma_1\} = T_0^{c-}(f_n) \rightarrow T_0^{c-}(f_0) \\
&= T_{t=0}^{c-}\{f_0(t) + \gamma_1\} = T_0^{c}(f_0 + \beta), \\
&| f_n(c) + \gamma_2 - f_n(c-) - \gamma_1 | + | f_n(c+) + \gamma_3 - f_n(c) - \gamma_2 | \\
&\quad \rightarrow | f_0(c) + \gamma_2 - f_0(c-) - \gamma_1 | + | f_0(c+) + \gamma_3 - f_0(c) - \gamma_2 |, \\
T_{c+}^{1}(f_n + \beta) &= T_{t=c+}^{1}\{f_n(t) + \gamma_3\} = T_{c+}^{1}(f_n) \rightarrow T_{c+}^{1}(f_0) \\
&= T_{t=c+}^{1}\{f_0(t) + \gamma_3\} = T_{c+}^{1}(f_0 + \beta).
\end{aligned}$$

Combining these three relations establishes the lemma.

LEMMA 6.2. *If β (in BV) is a singular function of the saltus type, then the relation $f_n - uv \rightarrow f_0$ implies the relation $(f_n+\beta) - uv \rightarrow (f_0+\beta)$.*

There exist elementary step-functions $\beta_1, \beta_2, \beta_3, \cdots$ such that $\beta(t) = \sum_{j=1}^{\infty}\beta_j(t)$ for t on $[0, 1]$ with $T_0^1(\beta) = \sum_{j=1}^{\infty} T_0^1(\beta_j)$. Hence $T_0^1(B_p - \beta) \rightarrow 0$, where $B_p(t) = \sum_{j=1}^{p}\beta_j(t)$. From the preceding lemma we conclude (by induction) that

$$\lim_{n\rightarrow\infty} \|f_n + B_p\| = \|f_0 + B_p\| \qquad (p = 1, 2, 3, \cdots).$$

Thus follows

$$\begin{aligned}
\|f_0 + \beta\| &\leqq \liminf_{n\rightarrow\infty} \|f_n + \beta\| \leqq \limsup_{n\rightarrow\infty} \|f_n + \beta\| \\
&\leqq \limsup_{n\rightarrow\infty} \|f_n + B_p\| + \|\beta - B_p\| = \|f_0 + B_p\| + \|\beta - B_p\| \\
&\leqq \|f_0 + \beta\| + 2\|\beta - B_p\| \rightarrow \|f_0 + \beta\| \quad \text{as} \quad p \rightarrow \infty,
\end{aligned}$$

and the proof is completed.

LEMMA 6.3. *If X_0 is an a.c. function in RBV, then the relations $X_n - l \rightarrow X_0$ and $Y_n - l \rightarrow Y_0$ imply the relation $(X_n + iY_n) - v \rightarrow (X_0 + iY_0)$.*

Define $\alpha(t) = Y_0(0) + \int_0^t Y'(s)ds$, $\beta(t) = Y(t) - \alpha(t)$ for t on $[0, 1]$ noting that β is singular with $\beta(0) = 0$. From Theorems 5.1 and 5.4 we have $\|Y_n - \alpha\| \rightarrow \|\beta\|$ and since $\|X_n - X_0\| \rightarrow 0$ we deduce $\|X_n + i\alpha\| \rightarrow \|X_0 + i\alpha\|$, so that

$$\begin{aligned}
\|X_n + iY_n\| &= \|X_n + i(Y_n - \alpha + \alpha)\| \\
&\leqq \|X_n + i\alpha\| + \|i(Y_n - \alpha)\| \rightarrow \|X_0 + i\alpha\| + \|i\beta\| \\
&= \|X_0 + i(\alpha + \beta)\| = \|X_0 + iY_0\|,
\end{aligned}$$

and the proof is complete (by semi-continuity).

THEOREM 6.2. *If X_0 is a pseudo-absolutely continuous function in RBV, then the relations $X_n - ul \rightarrow X_0$ and $Y_n - ul \rightarrow Y_0$ imply the relations*

$$(X_n + Y_n) - ul \rightarrow (X_0 + Y_0) \quad \textit{and} \quad X_nY_n - ul \rightarrow X_0Y_0.$$

Since X_0 is pseudo-absolutely continuous we conclude the existence of an a.c. function α in *RBV* and a singular function β of the saltus type likewise in *RBV* for which $X_0 = \alpha + \beta$. As a corollary of Lemma 6.2, $(X_n - \beta) - ul \rightarrow \alpha$, and since $Y_n - ul \rightarrow Y_0$ it is seen from Lemma 6.3 that

$$(X_n - \beta + iY_n) - uv \rightarrow (\alpha + iY_0),$$

so that using Lemma 6.2 again, yields

$$(X_n + iY_n) - uv \rightarrow (\alpha + \beta + iY_0) = (X_0 + iY_0).$$

Letting $\phi_1(x+iy)=x+y$ and $\phi_2(x+iy)=xy$ for $-\infty<x, y<\infty$ we conclude from Theorem 4.2 that

$$\phi_1:(X_n+iY_n)-uv\rightarrow\phi_1:(X_0+iY_0),$$
$$\phi_2:(X_n+iY_n)-uv\rightarrow\phi_2:(X_0+iY_0).$$

That is,

$$(X_n+Y_n)-uv\rightarrow(X_0+Y_0)\quad\text{and}\quad X_nY_n-uv\rightarrow X_0Y_0.$$

Application of Theorem 5.7 completes the proof.

7. **Strong convergence.** It is at once apparent that strong convergence implies every other type considered in this paper, and also that it is invariant under addition and multiplication. It is natural to ask if Theorem 6.1 likewise holds for strong convergence. The answer is yes, but before proving this we state as an obvious corollary of Lemma 4.5 the following

LEMMA 7.1. *If Y is a pseudo-absolutely continuous function in RBV and u is applicable (R) to Y, then $(u|Y)$ is likewise pseudo-absolutely continuous.*

We are now prepared to prove

THEOREM 7.1. *If u is applicable (R) to Y_n for $n=0, 1, 2, \cdots$, then the relation $Y_n-s\rightarrow Y_0$ implies*

$$(u|Y_n)-s\rightarrow(u|Y_0).$$

Since Y_0 is in RBV there exist a continuous function α in RBV and a singular function β of the saltus type likewise in RBV for which $Y_0=\alpha+\beta$. Defining

$$S(t)=\frac{t+T_0^t(\alpha)}{1+T_0^1(\alpha)}\qquad(0\leqq t\leqq 1),$$

it is seen that S is a continuous increasing function in RBV. Clearly there exists an increasing function Ψ satisfying a Lipschitz condition on $E_x[-\infty<x<\infty]$ for which $\Psi\{S(t)\}=t$, $0\leqq t\leqq 1$. Let $A(s)=\alpha\{\Psi(s)\}$, $B(s)=\beta\{\Psi(s)\}$, $\eta_n(s)=Y_n\{\Psi(s)\}$, $(0\leqq s\leqq 1;\ n=0, 1, 2, \cdots)$. Finally let $u_1(x, y)=u(\Psi(x), y)$, $-\infty<x, y<\infty$.

First notice that

$$\begin{aligned}|A(s_2)-A(s_1)|&\leqq T_{\Psi(s_1)}^{\Psi(s_2)}(\alpha)<\Psi(s_2)-\Psi(s_1)+T_{\Psi(s_1)}^{\Psi(s_2)}(\alpha)\\&=(1+T_0^1(\alpha))(S\{\Psi(s_2)\}-S\{\Psi(s_1)\})=(1+T_0^1(\alpha))(s_2-s_1)\end{aligned}$$

if $0\leqq s_1\leqq s_2\leqq 1$, which implies absolute continuity of A. Since β is a singular function of the saltus type, it may be seen from Lemma 3.8 that B is likewise

a singular function of the saltus type so that noting $\eta_0 = A + B$ establishes the pseudo-absolute continuity of η_0. Now from the definitions involved and the fact that Ψ is increasing and absolutely continuous it becomes apparent that if u is $\mathfrak{R}$, then u_1 is $\mathfrak{R}$; if u is $\mathfrak{R}_2$, then u_1 is $\mathfrak{R}_2$; and since monotonicity of Y_n implies monotonicity of η_n for $n = 0, 1, 2, \cdots$, we conclude that u_1 is applicable (R) to η_n for $n = 0, 1, 2, \cdots$. Another application of Lemma 3.8 yields the relation

$$\|\eta_n - \eta_0\| = \|Y_n - Y_0\| \to 0,$$

so that, since strong convergence implies uniform convergence in length, we may deduce successively (with the help of Lemma 7.1, Theorems 6.1 and 6.2) the following relations,

$$\eta_n - ul \to \eta_0, \quad (u_1|\eta_n) - ul \to (u_1|\eta_0), \quad \{(u_1|\eta_n) - (u_1|\eta_0)\} - ul \to \theta,$$
$$\|(u_1|\eta_n) - (u_1|\eta_0)\| \to 0.$$

Thus (Lemma 3.8)

$$\begin{aligned} T_{t=0}^{1}\{u(t, Y_n(t)) - u(t, Y_0(t))\} &= T_{s=0}^{1}\{u[\Psi(s), Y_n(\Psi(s))] - u[\Psi(s), Y_0(\Psi(s))]\} \\ &= T_{s=0}^{1}\{u(\Psi(s), \eta_n(s)) - u(\Psi(s), \eta_0(s))\} = T_{s=0}^{1}\{u_1(s, \eta_n(s)) - u_1(s, \eta_0(s))\} \\ &\leqq \|(u_1|\eta_n) - (u_1|\eta_0)\| \to 0, \end{aligned}$$

which implies immediately

$$\|(u|Y_n) - (u|Y_0)\| \to 0$$

as was to be proved.

THEOREM 7.2. *If f_0 is an a.c. function in BV, then a necessary and sufficient condition that $f_n - s \to f_0$ is that*

$$(cI + f_n) - v \to (cI + f_0),$$

for all real numbers c.

The necessity being obvious we turn to the sufficiency. Let $X_n + iY_n = f_n$ for $n = 0, 1, 2, \cdots$, and define

$$\phi_1(x + iy) = x, \ \phi_2(x + iy) = x + y, \text{ for } -\infty < x, \ y < \infty.$$

Since ϕ_1 and ϕ_2 are applicable to $(cI + f_n)$ for $n = 0, 1, 2, \cdots$ and strictly applicable to $(cI + f_0)$ whatever real number c may be, we conclude

$$(cI + X_n) - v \to (cI + X_0), \quad (cI + X_n + Y_n) - v \to (cI + X_0 + Y_0)$$

for all real numbers c. Whence, with the help of Theorems 5.5 and 5.4 follow successively the relations

$$X_n - l \to X_0, \quad (X_n + Y_n) - l \to (X_0 + Y_0), \quad Y_n - l \to Y_0,$$
$$\|X_n - X_0\| \to 0, \quad \|Y_n - Y_0\| \to 0, \quad \|f_n - f_0\| \to 0$$

and the proof is complete.

8. Applications; generalizations of a theorem of Plessner and its converse. It is well known that if $p>1$ then the two relations

$$\int_0^t f_n(s)ds \to \int_0^t f_0(s)ds \qquad (0 \leqq t \leqq 1)$$

and

$$\int_0^1 |f_n(t)|^p dt \to \int_0^1 |f_0(t)|^p dt$$

imply and are implied by the relation

$$\int_0^1 |f_n(t) - f_0(t)|^p dt \to 0.$$

It is likewise well known that the theorem is not true for $p=1$. The following theorem would therefore seem of some interest.

THEOREM 8.1. *If f_n is a summable function in CR for $n=0, 1, 2, \cdots$, then the two relations*

$$\int_0^t f_n(s)ds \to \int_0^t f_0(s)ds \qquad (0 \leqq t \leqq 1),$$

$$\int_0^1 |c + f_n(t)|\, dt \to \int_0^1 |c + f_0(t)|\, dt \text{ for all real numbers } c,$$

imply and are implied by the relation

$$\int_0^1 |f_n(t) - f_0(t)|\, dt \to 0.$$

Obviously the last relation implies the first two. Assuming the first two relations to be true and defining $F_n(t) = \int_0^t f_n(s)ds$ (for $0 \leqq t \leqq 1$; $n=0, 1, 2, \cdots$), we note from the first relation that $(ct+F_n(t)) \to (ct+F_0(t))$ for t on $[0, 1]$, and from the second relation that

$$T_{t=0}^{1}\{ct + F_n(t)\} = \int_0^1 |c + f_n(t)|\, dt \to \int_0^1 |c + f_0(t)|\, dt = T_{t=0}^{1}\{ct + F_0(t)\}$$

for all real numbers c. Hence

$$(cI + F_n) - v \to (cI + F_0) \text{ for all real numbers } c,$$

so that Theorem 7.2 implies

$$\int_0^1 | f_n(t) - f_0(t) | \, dt = T_0^1(F_n - F_0) \to 0$$

which completes the proof.

A result of Ursell combined with a theorem of Plessner establishes the following theorem:†

Let f be a finite, real-valued, measurable function with period 1. *If*

$$\lim_{h \to 0} T_{t=0}^1 \{ f(t + h) - f(t) \} = 0,$$

then f is a.c. on $[0, 1]$.

We now propose to generalize this theorem. First, however, it is convenient to prove the following

THEOREM 8.2. *Let f be a finite, real-valued, function with period* 1 *which is measurable on a set of positive measure. If there exists a non-vanishing function g in RBV such that*

$$\lim_{h \to 0} T_{t=0}^1 \{ f(t) + hg(t) \} = 0,$$

then f is continuous on $[0, 1]$.

Clearly there exists a closed set D on $[0, 1]$ of positive measure relative to which f is continuous. Let β be the characteristic function of this set and denote $\int_0^y \beta(x)dx$ by $B(y)$ for $-\infty < y < \infty$. Now B satisfies a Lipschitz condition so that use of Theorems 7.1 and 3.1, Lemma 3.7, and Corollary 3.1 establishes the relation

$$\begin{aligned} 0 &= \lim_{h \to 0} T_{t=0}^1 \{ B(t + hg(t)) - B(t) \} \\ &\geqq \limsup_{h \to 0} \int_0^1 | B'(t + hg(t))(1 + hg'(t)) - B'(t) | \, dt \\ &= \limsup_{h \to 0} \int_0^1 | \beta(t + hg(t))(1 + hg'(t)) - \beta(t) | \, dt. \end{aligned}$$

Hence there exists a $\delta_0 > 0$ such that $|h| < \delta_0$ implies

$$\int_0^1 | \beta(t + hg(t))(1 + hg'(t)) - \beta(t) | \, dt < \int_0^1 |\beta(t)| dt,$$

† An elementary proof was given by N. Dunford, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 356–358.

so that corresponding to each h^* for which $|h^*| < \delta_0$ there exists a point t^* in D, such that $t^* + h^*g(t^*)$ is likewise in D; for assuming the contrary leads immediately to a contradiction of the above relation.

Let ϵ be any positive number. Since D is closed there exists a $\delta_1 > 0$ such that, $|h| < \delta_1$ implies

$$|f(t + hg(t)) - f(t)| < \frac{\epsilon}{2}$$

for all t on $[0, 1]$ for which t and $t + hg(t)$ are both in D. By hypothesis there exists a $\delta_2 > 0$ such that $|h| < \delta_2$ implies

$$T_{t=0}^{1}\{f(t + hg(t)) - f(t)\} < \frac{\epsilon}{2}.$$

Let δ be the least of the numbers δ_0, δ_1, δ_2, and let h_0 be any number $< \delta$ in absolute value. As we have seen, there exists a point t_0 in D such that $t_0 + h_0g(t_0)$ is likewise in D. Hence

$$\begin{aligned} |f(t_1 + h_0g(t_1)) - f(t_1)| &\leqq |f(t_1 + h_0g(t_1)) - f(t_1) - f(t_0 + h_0g(t_0)) + f(t_0)| \\ &\quad + |f(t_0 + h_0g(t_0)) - f(t_0)| \\ &< T_{t=0}^{1}\{f(t + h_0g(t)) - f(t)\} + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for t_1 on $[0, 1]$. Hence f is continuous on $[0, 1]$ and the proof is complete.

It should be noted that the only place in the proof where a result of this paper is used is in proving

$$\lim_{h\to 0} \int_0^1 |\beta(t + hg(t))(1 + hg'(t)) - \beta(t)|\, dt = 0.$$

However if $g(t) = 1$ for t on $[0, 1]$, this relation is an immediate consequence of a well known theorem of Lebesgue, which in connection with the method used in proving Theorem 8.3 leads to a proof of the Plessner theorem which is independent of the preceding results in this paper.

We now turn to

THEOREM 8.3. *Let f be a finite, real-valued function with period* 1 *which is measurable on a set of positive measure. If there exist a function g in RBV and a positive number r such that*

$$\lim_{h\to 0} T_{t=0}^{1}\{f(t + hg(t)) - f(t)\} = 0$$

with $|g(t)| > r$ for t on $[0, 1]$, *then f is a.c. on* $[0, 1]$.

Let $\delta>0$ be such that $T_{t=0}^{1}U(t, h)<1$ for $|h|\leqq\delta$, where $U(t, h) = [f(t+hg(t))-f(t)]g(t)$ for $0\leqq t\leqq 1$, $h\leqq\delta$; let $\alpha(t)=1/g(t)$, $(0\leqq t\leqq 1)$; let $r_0=r^{-1}$ and note that $\|\alpha\|<\infty$, $|\alpha(t)|<r_0$ for t on $[0, 1]$; let $V(h)=T_{t=0}^{1}U(t, h)$ for $|h|\leqq\delta$. Since V is bounded on $[-\delta, \delta]$ and f is continuous by the previous theorem, the semi-continuity property of total variation shows that V is lower semi-continuous and hence summable on $[-\delta, \delta]$. Let

$$M(h) = \sup_{0\leqq t\leqq 1,\, |s|\leqq h} U(t, s) \qquad (0 \leqq h \leqq \delta)$$

remarking that M is monotone on $[0, \delta]$ and $M(h)\to 0$ as $h\to 0+$. Define

$$F_h(t) = \frac{1}{h}\int_t^{t+h} f(s)ds \qquad (0 \leqq t \leqq 1, 0 \leqq h \leqq r_1),$$

where $r_1=\delta r$, and let S be any partition of $[0, 1]$ with $S\equiv(0=t_0<t_1<t_2 < \cdots <t_N=1)$.

From the relation

$$\begin{aligned} F_h(t) - f(t) &= \frac{1}{h}\int_0^h \{f(t + s) - f(t)\}ds \\ &= \frac{1}{h}\int_0^{h\alpha(t)} \{f(t + sg(t)) - f(t)\}g(t)ds \\ &= \frac{1}{h}\int_0^{h\alpha(t)} U(t, s)ds \qquad (0 \leqq t \leqq 1, 0 \leqq h \leqq r_1) \end{aligned}$$

and the relation

$$\begin{aligned} &\sum_{j=1}^{N}\left|\frac{1}{h}\int_0^{h\alpha(t_j)} U(t_j, s)ds - \frac{1}{h}\int_0^{h\alpha(t_{j-1})} U(t_{j-1}, s)ds\right| \\ &= \sum_{j=1}^{N}\left|\frac{1}{h}\int_0^{h\alpha(t_j)} U(t_j, s) - U(t_{j-1}, s)ds + \frac{1}{h}\int_{h\alpha(t_{j-1})}^{h\alpha(t_j)} U(t_{j-1}, s)ds\right| \\ &\leqq \sum_{j=1}^{N}\left\{\frac{1}{h}\int_{-hr_0}^{hr_0} | U(t_j, s) - U(t_{j-1}, s) | ds + \frac{1}{h}\left|\int_{h\alpha(t_{j-1})}^{h\alpha(t_j)} M(hr_0)ds\right|\right\} \\ &\leqq \frac{1}{h}\int_{-hr_0}^{hr_0} \sum_{j=1}^{N} | U(t_j, s) - U(t_{j-1}, s) | ds + \sum_{j=1}^{N} M(hr_0) | \alpha(t_j) - \alpha(t_{j-1}) | \\ &\leqq \frac{1}{h}\int_{-hr_0}^{hr_0} V(s)ds + M(hr_0)\|\alpha\| \qquad (0 \leqq h \leqq r_1), \end{aligned}$$

we deduce (since S was arbitrary) the relation

$$T_0^1(F_h - f) \leqq \frac{1}{h}\int_{-hr_0}^{hr_0} V(s)ds + M(hr_0)\|\alpha\| \qquad (0 \leqq h \leqq r_1).$$

Now, Lemma 3.2 yields the relation

$$V(s) \leqq [\,|\, f(sg(0)) - f(0)\,| + T_{t=0}^{1}\{f(t + sg(t)) - f(t)\}\,]\|g\| \qquad (\,|\, s \,|\, \leqq r_1),$$

so that $V(s) \to 0$ as $s \to 0$. Hence

$$\lim_{h \to 0+} T_0^1\{F_h - f\} = 0,$$

where F_h is a.c. for $0 \leqq h \leqq r_1$ which implies, as is well known, the absolute continuity of f on $[0, 1]$. This completes the proof.

From the results in §7 it is clear that a variety of theorems concerning the behavior of $T_{t=0}^1\{f(t+hg(t)) - f(t)\}$ as $h \to 0$ (or as $h \to 0+$, $h \to 0-$) can be readily proved. Among these is one which can be proved directly without great difficulty, and which forms the necessity part of the next and concluding theorem. This theorem is a simultaneous extension of Plessner's theorem and its converse.

THEOREM 8.4. *Let f be a finite, real-valued function with period* 1 *which is measurable on set of positive measure. Let g be a non-vanishing function in RBV which satisfies a Lipschitz condition. Then a necessary and sufficient condition that f be a.c. on* $[0, 1]$ *is that*

$$\lim_{h \to 0} T_{t=0}^{1}\{f(t + hg(t)) - f(t)\} = 0.$$

Simply note that $t + hg(t)$ increases with t for h sufficiently small and apply Theorem 7.1.

BROWN UNIVERSITY,
PROVIDENCE, R. I.

GENERALIZED DERIVATIVES AND APPROXIMATION BY POLYNOMIALS*

BY

W. E. SEWELL

1. **Introduction.**† Let E be a closed limited point set in the complex z-plane, let the complement (with respect to the extended plane) K of E be simply connected, and let C denote the boundary of E. In the case where E contains interior points we shall denote by R the limited simply connected region whose closure is $\bar{R} \equiv E$ and whose boundary is C. Walsh [1]‡ has shown that if R is a Jordan region and if the function $f(z)$ is analytic in R and continuous in $\bar{R}$, then $f(z)$ can be uniformly approximated in $\bar{R}$ by a polynomial in z. The best degree of convergence of approximating polynomials to the function depends upon the continuity properties both of the boundary of the region and of the function on the boundary. Reciprocally, under certain conditions continuity properties of the function on the boundary are a consequence of the degree of convergence of approximating polynomials. *It is the purpose of this paper to investigate the nature of the relation between continuity properties and degree of convergence.*

This investigation can be divided into two parts which we shall call *Problem α* and *Problem β*. In *Problem α* either the continuity properties of the function $f(z)$, analytic in R, are given in the closed limited simply connected region $\bar{R}$, bounded by a Jordan curve C, and we study the degree of convergence of certain approximating polynomials in $\bar{R}$, or the degree of convergence in $\bar{R}$ is given and we study the continuity properties of the function on the boundary C. *In Problem α then we study degree of convergence where the function is not analytic in the closed region.*

In *Problem β* we study the degree of convergence of approximating polynomials on interior sets. Let $w=\phi(z)$ map the exterior of E conformally on the exterior of the unit circle, $|w|=1$, so that the points at ∞ correspond to each other, then the image in the z-plane of the circle $|w|=\rho$, $\rho>1$, shall be designated by C_ρ (as above C will denote the boundary of E). If the function $f(z)$ is analytic interior to a particular C_ρ and continuous in $\bar{C}_\rho$, the closed

* Presented to the Society, December 27, 1934; received by the editors February 29, 1936, and, in revised form, July 21, 1936.

† I wish to express my thanks to Professor J. L. Walsh who proposed this problem to me and under whose direction this paper was written as a thesis at Harvard University. I am also indebted to Dr. W. Seidel who, during Professor Walsh's sabbatical leave, gave me many valuable suggestions.

‡ The number in brackets refers to the bibliography at the end of this paper.

limited simply connected region bounded by C_ρ, the problem here is *the relation between the degree of convergence of approximating polynomials on E and continuity properties of the function* $f(z)$ *on* $\overline{C}_\rho$.

Problem α for E a segment of the axis of reals has been widely studied.* Here we follow Montel [1] in describing the continuity properties of the function by the order of its generalized derivative.† We also use to advantage the theory of approximation to functions analytic in closed regions.‡

In studying degree of convergence we need a close evaluation of the *maximum and minimum distances from the boundary C to the curve* C_ρ, *which we shall call* $D(C, C_\rho)$ *and* $d(C, C_\rho)$, *respectively*. Chapter I is devoted to an investigation of these distance functions for various types of boundaries. We find *for C a Jordan curve with corners that* $D(C, C_\rho) \leqq M_2(\rho-1)^s$, $d(C, C_\rho) \geqq M_1(\rho-1)^t$, *where* $0<s\leqq 1$, $1\leqq t<2$. *For E an arbitrary closed limited point set whose complement is simply connected we show that* $d(C, C_\rho) \geqq M_1(\rho-1)^2$.

Due to the fact that the generalized derivatives are defined by improper integrals evaluated along rectifiable curves, and to the fact that our method demands the absolute convergence of these integrals, Chapter II is taken up with a discussion, principally by examples, of the types of *curves along which the integral* $\int_k^z |z-x|^{-\alpha}|dx|$, $0<\alpha<1$, *converges*. We show, for example, that *there are rectifiable curves for which this integral diverges*.

In Chapter III the various properties of the generalized derivative which we need for application to approximation are investigated. Also for certain curves relations are established between Lipschitz conditions and generalized derivatives.

Theorems of Bernstein [2], Riesz [1], Markoff [1], and Montel [1] on derivatives of polynomials are extended to generalized derivatives for various types of regions in Chapter IV. In considering Jordan curves with corners we obtain a generalization of a theorem of Szegö [2]. We show that *if* $P_n(z)$ *is a polynomial of degree n in z and* $|P_n(z)| \leqq M$ *on C, then for C an analytic Jordan curve* $|P_n^\alpha(z)|$§ $\leqq MM_1(\alpha, C)n^\alpha$, *for C a curve with corners* $|P_n^\alpha(z)| \leqq MM_1(\alpha, C)n^{\alpha t}$, $t<2$, *and for C the boundary of a limited simply connected region every boundary point of which is accessible* (see §23)

$$|P_n^\alpha(z)| \leqq MM_1(\alpha, C)n^{2\alpha},$$

where in all cases $\alpha>0$.

* See e.g., Bernstein [2], Jackson [1], Montel [1].

† Liouville [1], Riemann [1]. See also Weyl [1], Levy [1], Hardy and Littlewood [1, 2], Tamarkin [1], Doetsch [1]. For further references see the articles by Hardy and Littlewood, and Tamarkin.

‡ For a complete and excellent exposition of these results see Walsh [1].

§ $f^\alpha(z)$ denotes the generalized derivative of order α of $f(z)$.

In Chapter V we apply the results of the preceding chapters to obtain theorems on approximation by polynomials leading to solutions of various cases of *Problem* α. For example, *if* $f(z)$ *is analytic in* R, *a limited simply connected region every boundary point of which is accessible* (see §23), *and continuous in* $\overline{R}$, *and if for every* n *there exists a polynomial* $P_n(z)$ *of degree* n *in* z *such that* $|f(z)-P_n(z)|\leqq M/n^\alpha$, z *in* $\overline{R}$, *then* $f(z)$ *has a bounded generalized derivative of order* $\alpha'<\alpha/t$ *on* C, *the boundary of* R, *where* $2\geqq t\geqq 1$, *depending on the continuity properties of* C. It should be noted here that the analyticity of $f(z)$ in R, its continuity in $\overline{R}$, and the continuity of an ordinary derivative of $f(z)$ on C imply the continuity of this derivative of $f(z)$ in the closed region bounded by C;* thus it is sufficient in establishing the results of this paper on approximation to assume the continuity merely of the function $f(z)$ in the closed region.

The last chapter is devoted to a study of *Problem* β. Here we consider *uniformly bounded functions as well as functions continuous in closed regions.* As far as the relation between degree of convergence of a sequence $P_n(z)$ to $f(z)$ on E and the continuity properties of $f(z)$ on C_ρ we prove, for instance, that *if* E *is a closed limited point set whose complement is simply connected, and if for every* n *there exists a polynomial* $P_n(z)$ *of degree* n *in* z *such that* $|f(z)-P_n(z)|\leqq M/(n^{\alpha+1}\rho^n)$, $\alpha>0$, $\rho>1$, z *on* E, *then* $f(z)$ *has derivatives of all orders* $\alpha'<\alpha$ *on* C_ρ, *and show by example that this is the best result possible in a certain sense.*

CHAPTER I

THE LEVEL CURVES OF THE GREEN'S FUNCTION

2. **Definitions.** Let E, K, C, and C_ρ have the same meaning as above. The curve C_ρ is a level curve of the Green's function of K with pole at ∞ and as ρ approaches 1 the analytic Jordan curve C_ρ approaches the boundary C of E, however, if C is not a Jordan curve there may be points of C which C_ρ does not approach. *Let* P *be a point of* C *and define* $d(P, C_\rho)$ *as the greatest lower bound of the distances from* P *to the points of* C_ρ. Now *we define* $d(C, C_\rho)$ *as the greatest lower bound of* $d(P, C_\rho)$ *as* P *traverses* C, *and* $D(C, C_\rho)$ *as the least upper bound of* $d(P, C_\rho)$ *as* P *traverses* C. In a similar manner we define $d(C_\rho, C)$ and $D(C_\rho, C)$. We will investigate here the nature of the approach of C_ρ to C by studying the functions d and D. In the case where C is a Jordan curve all of these functions approach 0 as ρ approaches 1; for an arbitrary boundary this is not necessarily true. Although C_ρ is defined for every $\rho>1$ we are interested in the behavior of C_ρ for ρ near 1, or at least for ρ uniformly bounded from infinity and hence this condition will be assumed in all our inequalities.

* For a detailed discussion of this fact see Walsh and Sewell [1].

In the above discussion, we have assumed that the set E is limited, we may consider the case, however, where R is an arbitrary simply connected region, containing, for definiteness, the point $z=0$, and denote by C the boundary of R. Let $w=\phi(z)$ map conformally the interior of R on the interior of the unit circle, $|w|=1$, so that the origins correspond to each other, and consider the interior level curve C_ρ which is the image in the z-plane, under the inverse map, of the circle $|w|=\rho$, $0<\rho<1$. We can define the functions d and D precisely as above and in general the results are valid for both interior and exterior level curves. Of course, in the case where $\overline{R}$ contains the point at ∞ we can obtain no evaluation for $D(C, C_\rho)$.

3. **Smooth curves.** Let C be an analytic Jordan curve and let $w=\phi(z)$, whose inverse is $z=\psi(w)$, map the interior of C on the interior of $|w|=1$. Then we know that*

$$0 < N_1 < |\psi'(w)| < N_2 < \infty, \quad \text{for} \quad |w| \leqq 1, \tag{3.10}$$

and hence by considering the difference quotient we have

$$d(C, C_\rho) \geqq M_1 |1 - \rho|, \tag{3.11}$$

$$D(C, C_\rho) \leqq M_2 |1 - \rho|, \tag{3.12}$$

where M_1 and M_2 are constants depending only on C. In the above evaluation we have used the analyticity of C only to establish inequality (3.10) on the first derivative of the mapping function in the closed unit circle. As a matter of fact inequality (3.11) and inequality (3.12) are valid in case the mapping function satisfies the inequality

$$\left| \frac{\psi(w_1) - \psi(w_2)}{w_1 - w_2} \right| \geqq N_1 > 0, \tag{3.13}$$

and the inequality

$$\left| \frac{\psi(w_1) - \psi(w_2)}{w_1 - w_2} \right| \leqq N_2 < \infty, \tag{3.14}$$

respectively, both holding uniformly for $|w_1| \leqq 1$, $|w_2| \leqq 1$.

Definition 3.1. *If inequalities* (3.13) *and* (3.14) *are satisfied, we shall say that "C is a curve of Type S."*

The geometric properties of a curve of Type S are of interest. Seidel [1]† has shown that the mapping function possesses the property (3.10), and hence satisfies the inequalities (3.13) and (3.14), provided

* $f'(z)$ denotes the first derivative of $f(z)$.

† See in particular pp. 213–220. See also Carathéodory [1], Chap. VI; Visser [1]; Kellogg [1]; Warschawski [1].

$$(3.15) \qquad |\omega(s+h) - \omega(s)| < M|h|, \qquad M > 0,$$

where ω denotes the angle between the positively directed tangent and the real axis and s the arc length along the curve.

The same evaluations, (3.11) and (3.12), hold for $d(C_\rho, C)$ and $D(C_\rho, C)$. If suitably interpreted our definition and evaluations are valid for exterior level curves. Thus we have the following theorem:

THEOREM 3.2. *Let C be a curve of Type S. Then we have*

$$d(C, C_\rho) \geqq M_1|1-\rho|, \qquad D(C, C_\rho) \leqq M_2|1-\rho|,$$

where C_ρ is an exterior or interior level curve, and M_1 and M_2 are constants depending only on C.

4. **Curves with corners.** Let C be a Jordan curve consisting of a finite number of analytic Jordan arcs meeting in corners with exterior openings $\mu_k\pi$, $0<\mu_k\leqq 2$, $k=1, 2, \cdots, p$; let $w=\phi(z)$, whose inverse is $z=\psi(w)$, denote the exterior mapping function, and consider the exterior level curves. If P is a point of C, *not a corner*, condition (3.10) is satisfied in a neighborhood of P due to the analyticity of the arc containing P, and hence we have inequalities corresponding to (3.11) and (3.12) for P, however the constants M_1 and M_2 are functions of P. In order to obtain a uniform evaluation we must investigate the behavior of C_ρ near a corner. Let z_0 be a corner with exterior opening $\mu\pi$, $0<\mu\leqq 2$, and consider the quotient

$$(4.10) \qquad \theta(z_1, z_2) = \frac{[\phi(z_1) - \phi(z_2)]^t}{z_1 - z_2},$$

where $t=\mu$ if $\mu\geqq 1$, and $t=1$ if $\mu<1$, z_1 and z_2 in the closed exterior of C. We can show that $|\theta(z_1, z_2)|$ is uniformly bounded by using the fact that in a sufficiently small neighborhood of a point not a corner the difference quotient of the mapping function is bounded in modulus due to the analyticity of the arc, and in the neighborhood of a corner the function $\phi(z)$ can be written in the form*

$$(4.11) \qquad \phi(z) - \phi(z_0) = (z - z_0)^{1/\mu}\lambda(z),$$

where $\lambda(z)$ is analytic in a neighborhood of the corner z_0 lying outside of C, is continuous in the closed neighborhood, and $\lambda(z_0)\neq 0$. This gives us an estimate for both $d(C, C_\rho)$ and $D(C, C_\rho)$.

In fact it is not necessary that the arcs be analytic so long as the difference quotient of the mapping function is bounded in modulus on every proper

* Osgood and Taylor [1], pp. 282–283. See also Warschawski [2], p. 324.

sub-arc, not containing an end point, of the arc joining two adjacent corners, and on this basis we will formulate the following definitions:

DEFINITION 4.1. *Let C be a Jordan curve composed of a finite number of Jordan arcs meeting in corners $z_1, z_2, \cdots, z_p$, of exterior openings $\mu_1\pi, \mu_2\pi, \cdots, \mu_p\pi$, $2>\mu_1\geqq\mu_2\geqq \cdots \geqq\mu_p>0$, and let the difference quotient of the mapping function $w=\phi(z)$, which maps the exterior of C on the exterior of $|w|=1$ so that the points at ∞ correspond to each other, be bounded in modulus on each smooth sub-arc. Let $t=\mu_1$ if $\mu_1\geqq 1$, and $t=1$ if $\mu_1<1$. Then we shall say that "C is a curve of Type t."*

Attention should be called to the fact that in Definition 4.1 we have excluded an exterior opening of 2π, i.e., $\mu_1=2$. This is purely for convenience in later application; *in fact all of the results of this paragraph are valid for $\mu_1=2$.*

DEFINITION 4.2. *Under the same conditions as in Definition 4.1 let $s=\mu_p$ if $\mu_p\leqq 1$, and $s=1$ if $\mu_p>1$. Then we shall say that "C is a curve of Type s."*

And now we can state the following theorems:

THEOREM 4.3. *Let C be a curve of Type t. Then we have*

$$d(C, C_\rho) \geqq M_1(\rho - 1)^t, \qquad \rho > 1, \qquad 1 \leqq t < 2,$$

where C_ρ is an exterior level curve, and M_1 is a constant depending only on C.

THEOREM 4.4. *Let C be a curve of Type s. Then we have*

$$D(C, C_\rho) \leqq M_2(\rho - 1)^s, \qquad \rho > 1, \qquad 0 < s \leqq 1,$$

where C_ρ is an exterior level curve, and M_2 is a constant depending only on C.

These results hold for interior as well as exterior level curves if the interpretation of $\mu_1, \mu_2, \cdots, \mu_p$ is suitably modified.

It is to be noted here that we have excluded the case where *C has an exterior opening of zero.* At such points the function D is large, in fact Szegö [2, pp. 57–59] has shown that a Jordan curve can be constructed with a zero opening for which $D(C, C_\rho)$ decreases at any preassigned rate, however slow.

5. **More general regions.** In general, if

$$\left|\frac{\psi(w_1) - \psi(w_2)}{(w_1 - w_2)^\alpha}\right|, \qquad \alpha > 0, \qquad |w_1| \geqq 1, \qquad |w_2| \geqq 1, \tag{5.10}$$

considering the exterior map, is uniformly bounded we have $D(C, C_\rho) \leqq M_2(\rho-1)^\alpha$. Under a similar hypothesis on the uniform boundedness of (5.10) from zero we have $d(C, C_\rho) \geqq M_1(\rho-1)^\alpha$.

Moreover, if the point z_0 of C is mapped into a point w_0 of $|w|=1$, and the radius to w_0 cuts the circle $|w|=\rho$ in the point w_1, whose image in the z-plane is the point z_1 on C_ρ, we have

$$(5.11) \qquad d(z_0, C_\rho) \leqq \int_{z_1}^{z_0} |dz| = \int_{w_1}^{w_0} |\psi'(w)|\,|dw|$$

by using the Schwarz inequality* we have

$$(5.12) \qquad d(z_0, C_\rho) \leqq (\rho - 1)^{1/2}\left(\int_\rho^1 |\psi'(w)|^2 dr\right)^{1/2}.$$

Consequently *for a region with a boundary such that the square of the derivative of the mapping function is integrable absolutely along each radius, and the integral is uniformly bounded for the entire circumference, we have*

$$(5.13) \qquad D(C, C_\rho) \leqq M_2(\rho - 1)^{1/2}.$$

For the arbitrary simply connected region it is impossible to find a universal function of $\rho-1$, say $m(\rho-1)$, which approaches 0 with $\rho-1$, and such that $D(C, C_\rho) \leqq M_2 m(\rho-1)$. In fact, given an arbitrary function of $\rho-1$ with the property mentioned, Szegö [2] has constructed a region (actually a Jordan region) for which $D(C, C_\rho)$ decreases at precisely this rate.

6. **The function $d(C, C_\rho)$ for an arbitrary simply connected region.** We need the following lemma which follows directly from a result of Szegö [3]:

LEMMA 6.1. *Let R be a simply connected region in the z-plane containing the point $z=0$. Let $z=\psi(w)$ map R conformally on $|w|<1$ so that the origins correspond to each other. Then we have*†

$$\varliminf_{w\to e^{i\theta}} |\psi(w) - \psi(\alpha)| \geqq \frac{a}{16}(1-\rho)^2,$$

where $0<\rho<1$, $|\alpha|=\rho$, *and* $|\psi'(0)|=a$.

An immediate consequence of Lemma 6.1 is

THEOREM 6.2. *Let R be an arbitrary simply connected region with boundary C. Then for the interior level curves we have*

$$d(C, C_\rho) \geqq M_1(1-\rho)^2, \qquad 0<\rho<1,$$

where M_1 depends only on R.‡

* For this method see Ahlfors [1], pp. 7–8.

† $\varliminf$ denotes the lower limit and is used here since $\psi(e^{i\theta})$ may not be well defined.

‡ Professor Ahlfors has called my attention to the fact that this result contains, in a sense, the following theorem of Lindelöf [1] (see e.g., Walsh [1, pp. 27–32]):

For the exterior level curves we have

THEOREM 6.3. *Let E, with boundary C, be an arbitrary closed limited point set whose complement is simply connected. Then for the exterior level curves we have*

$$d(C, C_\rho) \geqq M_1(\rho - 1)^2, \qquad \rho > 1,$$

*where M_1 depends only on E.**

7. **Relation between ρ and the capacity.** Let E be an arbitrary closed limited point set in the z-plane whose complement K is simply connected. Let $z=\psi(w)$ map K conformally on $|w|>1$ so that the points at ∞ correspond to each other. Then

$$z = \psi(w) = cw + c_0 + \frac{c_1}{w} + \frac{c_2}{w^2} + \cdots, \quad c \neq 0; \tag{7.10}$$

by a rotation c can be made real and positive and we will assume this to be the case. *This constant c* is known as the *capacity*, or *Robin's constant*, or the *transfinite diameter*, or the *outer radius of the set E or of the boundary C of E*. We shall use the term *capacity* in this paper. If we denote by r the capacity of C_ρ, $\rho>1$, we have

$$r = \rho c, \dagger \tag{7.11}$$

a relation which will be found useful in our later work.

THEOREM. *Let R be a limited simply connected region in the z-plane containing the origin. Let Δ be the greatest diameter of R, and let δ be the shortest distance from $z=0$ to the boundary C of R. When R is mapped conformally on the interior of $|w|=1$ so that the origins in the two planes correspond to each other, every point of R at a distance less than r less than δ from C corresponds to a point w whose distance from $|w|=1$ is less than*

$$s(r) = \frac{2\log\Delta - 2\log\delta}{2\log\Delta - \log(\delta r)}.$$

Let a be an arbitrary point of C, draw a circle of radius r about a and consider a particular C_ρ which touches this circle. For this C_ρ we have $d(C, C_\rho)\leqq r$, but by Theorem 6.2 we know that $d(C, C_\rho)\geqq M_1(1-\rho)^2$. Thus we have $\rho \leqq r^{1/2}/M_1$, and since ρ is precisely $s(r)$ we obtain a much better upper bound for $s(r)$.

Also we can show by Lindelöf's method and the above result that any Jordan arc in the circle $|z-a|<r$ subtends an angle at $w=0$ which is in magnitude less than

$$4\sin^{-1}\left[\left(\frac{M_1}{2}\right)^{1/2}\Delta^{1/8}r^{1/8}\right].$$

* Professor Ahlfors has pointed out to me that this result holds for any set E whose complement K is connected and regular in the sense that K possesses a Green's function with pole at infinity. For a discussion of such sets see Walsh [1], pp. 65 ff.

† See e.g., Walsh [1], pp. 74–75.

Chapter II

Improper integrals along rectifiable curves

8. **Introduction and definition of curves of Type W.** In this chapter we will investigate the behavior, convergence, and divergence in particular, of some improper integrals evaluated along rectifiable Jordan arcs or curves, especially of the integral

$$(8.10) \qquad \int_k^z |z - x|^{-\alpha} |dx|, \qquad 0 < \alpha < 1,$$

where the path of integration is an arbitrary rectifiable Jordan arc or curve. The following definition serves as a basis for this investigation:

Definition 8.1. *A rectifiable Jordan arc or curve C shall be said to be an "arc or curve of Type W" if the integral* (8.10) *, where the path of integration is along C and k is an arbitrary but fixed point on C, converges uniformly in z, z on C. For $z=k$ we define $\int_k^z|z-x|^{-\alpha}|dx|=0$, on a curve or arc of Type W.*

9. **Properties of curves of Type W.** Using the above definition we obtain the following properties:

(9.1) *If on a rectifiable Jordan arc or curve C the chord and arc are infinitesimals of the same order uniformly, C is of Type W.*

This follows from the fact that we can replace the curvilinear integral by a rectilinear integral in the neighborhood of the point z.

(9.2) *Rectifiable Jordan arcs or curves with continuously turning tangents are of Type W.*

This is an immediate consequence of (9.1). Thus analytic Jordan arcs or curves are of Type W.

(9.3) *A curve of Type S is of Type W.*

This follows from the fact that (8.10) converges uniformly in the case where C is a circle.

(9.4) *Curves of Type t and curves of Type s are of Type W.*

This is an immediate consequence of (9.3), (9.1), and the fact that one side of a triangle is of the same infinitesimal order as the sum of the other two.

(9.5) *Not all rectifiable Jordan arcs and curves are of Type W.*

The following example illustrates this proposition:

Let $z=x+iy=(0, 0)$ and $k=(1, 0)$, and consider the arc formed by the *oblique* (*non-horizontal*) *sides* of the triangles whose vertices are

$$\left(\frac{1}{n}, 0\right), \quad \left(\frac{1}{n+1}, 0\right), \quad \text{and} \quad \left(\frac{2n+1}{2n(n+1)}, \frac{1}{n^{3/2}}\right), \qquad n = 1, 2, \cdots.$$

(9.6) *There are curves of Type W whose chord and arc are not infinitesimals of the same order.*

An example illustrating this is obtained by replacing $1/n^{3/2}$ in the above example by $\log n/n^2$. It is interesting to note here that if in this example we replace $1/n^{3/2}$ by $1/n^\beta$ the integral (8.10) converges for $\alpha<\beta-1$, the curve is rectifiable for $\beta>1$, and the chord and arc are not infinitesimals of the same order for $\beta<2$.

10. **Curves and regions of Type** W'. An arc or curve of Type W' is defined as follows:

DEFINITION 10.1. *Let C be an arc or curve of Type W. Let z_1 and z_2 be arbitrary points, $z_1 \neq z_2$, on C. Let z_3 be a point of C distinct from z_1 and z_2, and let z_2 lie between* z_1 and z_3. If constants M_1 and M_2, independent of z_1, z_2, and z_3, exist such that*

$$\int_{z_1}^{z_2} \left| z_3 - x \right|^{-\beta} \left| dx \right| \leqq M_1 \left| z_2 - z_3 \right|^{-\beta+1} + M_2, \qquad \beta > 0, \tag{10.10}$$

for z_2 arbitrarily near z_3, we shall say that C is an arc or curve of Type W'.

In case $\beta>1$ the constant M_2 may be taken as zero since the first term of the right-hand side of the expression (10.10) becomes infinite as the point z_2 approaches z_3.

A line segment is an arc of Type W'. Thus it follows that *any rectifiable Jordan arc or curve of which the chord and arc are infinitesimals of the same order uniformly is of Type W'.* Consequently *curves of Type S, Type t, and Type s are of Type W'.*

In connection with curves of Type W' we shall introduce a definition which will be found useful later.

DEFINITION 10.2. *Let R be a limited simply connected region in the z-plane. Let k be an arbitrary but fixed point of $\overline{R}$. If k can be joined to an arbitrary boundary point ζ of R by an arc γ of Type W' lying in $\overline{R}$, and if there exist constants M_1 and M_2 independent of ζ such that*

* Since C is a Jordan arc or curve the concept of "between" is well defined by fixing a direction on C; we may use the parametric representation with a preassigned direction on the line segment or the circle.

$$\int_k^z |\zeta - x|^{-\beta} |dx| \leqq M_1 |z - \zeta|^{-\beta+1} + M_2, \qquad \beta > 0, \tag{10.11}$$

for points z on γ and arbitrarily near ζ, we shall say that R is a region of Type W′.

The definition of a region of Type W is obvious and will not be stated explicitly.

The existence of a *single point k* in a region of Type W or W' is sufficient; from this we can show that any two points of $\overline{R}$ can be joined by an arc of Type W or W'.

Also it is clear that inequality (10.11) need be assumed only for z near the boundary, since by means, for example, of an interior level curve we can separate off a limited region bounded by an *analytic Jordan curve.*

Chapter III

Generalized derivatives

11. **Definition of the generalized derivative.** Let C be a rectifiable Jordan arc or curve in the z-plane and let $f(z)$ be continuous on C. The ordinary derivative, if it exists, of the function $f(z)$ at a point z_0 of C is defined as follows:

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \equiv \frac{df}{dz}\Big]_{z=z_0} \equiv f'(z_0), \tag{11.10}$$

z approaching z_0 along C. The generalized derivative, if it exists, of $f(z)$ at a point z_0 of C is defined as follows:

$$\begin{aligned} D_z^0 f(z_0) &\equiv f(z_0); \\ D_z^\alpha f(z_0) &\equiv \frac{1}{\Gamma(-\alpha)} \int_k^{z_0} (z_0 - x)^{-\alpha-1} f(x) dx, \qquad \alpha < 0; \\ D_z^\alpha f(z_0) &\equiv \frac{d^p}{dz^p} D_z^{\alpha-p} f(z)\Big]_{z=z_0}, \qquad 0 \leqq p - 1 \leqq \alpha < p, \end{aligned} \tag{11.11}$$

where p is a positive integer. Here the point k is an arbitrary but fixed point on C if C is a curve, and if C is an arc, k is one of the end points. The path of integration is along C in a fixed direction. The function $D_z^\alpha f(z)$, $\alpha<0$, is a function of z defined on C and the ordinary derivatives of such a function are defined by (11.10).

If α is a positive integer $D_z^\alpha f(z)$ reduces to an ordinary derivative of $f(z)$, since the exponent of $(z-x)$ in the integrand vanishes. If α is a negative integer, the function $D_z^\alpha f(z)$ is the integral of a single valued continuous function evaluated along a rectifiable Jordan arc or curve.

In the case where α is not an integer the function $(z-x)^{-\alpha-1}$ is multiple valued and care must be taken to choose a branch which varies continuously with z varying along C. First let us consider the case where C is a rectifiable Jordan arc and k one of its end points. We will cut the plane from k to ∞ along a path whose only point in common with C is k and which goes out to ∞ along the positive real axis; this cut makes the plane a simply connected region. Let $\beta=-\alpha-1$; by definition

$$(11.12)\qquad (z-k)^\beta = e^{\beta\log(z-k)} = e^{\beta\log|z-k|+i\beta\arg(z-k)},$$

and the various branches of this function arise in the arg $(z-k)$. The function $(z-k)^\beta$ is a multiple valued function of z defined throughout the z-plane except perhaps at the point $z=k$, where it can be defined by continuity as *zero* or *infinity*. We will define arg $(z-k)$ so that its limit as z approaches ∞ along the negative real axis is π. Thus we have in the simply connected region (cut plane) a well defined single valued analytic function of z. Now keep z fixed at an arbitrary point on C, $z\neq k$, and define arg $(z-x)$ for x on C, between k and z, in such a way that it takes the value arg $(z-k)$ determined above for $x=k$. If we hold z fixed on C the function $(z-x)^\beta$ for x varying along C from k to z is a single valued analytic function of x.

Now consider the case where C is a rectifiable Jordan curve and k is an arbitrary but fixed point on C. Cut the z-plane from k to ∞ along a path whose only point in common with $\overline{C}$, *the closed limited region bounded by* C, is the point k, and which goes out to ∞ along the positive real axis. Due to the fact that k is a branch point we exclude the value $z=k$ from consideration, in fact for a reason which will be apparent later we will exclude z from a certain preassigned two-dimensional neighborhood of k, for any fixed k. We will define the branch of $(z-k)^\beta$ precisely as above. For $(z-x)^\beta$, with z fixed on C, we will define the branch as above not only for x on C but also for x interior to C. It should be noted that x cannot make a loop around z since z is a boundary point and x is restricted to lie on or interior to C.

For the most general case let R be a region of Type W or W', let k be an arbitrary point of $\overline{R}$, and cut the plane from k to ∞ along a path which goes to ∞ along the positive real axis. This cut may contain points of $\overline{R}$ other than k, in any event we will exclude z from a two-dimensional strip containing the cut, and restrict z to lie on the boundary of R. We will restrict x to paths from k to z which contain no points of the cut. Under these restrictions our definition of the branch as defined above is valid.

It is not necessary to restrict z to the boundary of R. We can let z be any point of $\overline{R}$, in fact any point of the plane so long as we restrict x to paths from k to z which do not loop around z.

Thus we have in all cases a well defined single valued branch of the function $(z-x)^\beta$ which varies continuously with z under the restrictions prescribed above.

For a given rectifiable Jordan curve C the value of the derivative depends upon the choice of k, however, the existence of the derivative is independent of this choice. In fact we have

$$\int_k^z (z-x)^\beta f(x)dx = \int_k^{k_0} (z-x)^\beta f(x)dx + \int_{k_0}^z (z-x)^\beta f(x)dx,$$

or

$$\int_k^z (z-x)^\beta f(x)dx = \int_k^z (z-x)^\beta f(x)dx - \int_k^{k_0} (z-x)^\beta f(x)dx, \qquad k \neq k_0 \neq z. \tag{11.13}$$

We can evaluate the derivative of the second integral directly since it is a proper integral, and thus we see that *if a function has a bounded generalized derivative at a point z for a particular $k\neq z$, then it has a bounded generalized derivative of the same order for any k_0 between k and z.* The expression (11.13) indicates the type of function which represents the difference in value of the generalized derivative for two distinct choices of k.

Let $f(z)$ be analytic in a region R of Type W and continuous in $\overline{R}$. Let k be a fixed point of $\overline{R}$, and let C_1 and C_2 be two distinct arcs of Type W joining k to a boundary point z of R, and lying entirely in $\overline{R}$. We assume here that $z\neq k$ and that z does not lie on the cut; we also assume that C_1 and C_2 have no points in common with the cut except the point k, and, for simplicity, that C_1 and C_2 intersect only at k and z. If we draw a circle γ about z of radius r sufficiently small and let Δ be the region bounded by C_1, C_2, and one or more arcs of γ, we have by Cauchy's integral theorem

$$\int_\delta (z-x)^\beta f(x)dx = 0,$$

where δ denotes the boundary of Δ. Since $f(x)$ is continuous and $\beta>-1$ the modulus of this integral over the arc or arcs of γ under consideration approaches 0 as r approaches 0. Also the integrals along C_1 and C_2 converge as r approaches 0 since C_1 and C_2 are of Type W. Thus we have

$$\int_{C_1} (z-x)^\beta f(x)dx = \int_{C_2} (z-x)^\beta f(x)dx, \qquad \beta > -1. \tag{11.14}$$

Consequently the value of the integral is independent of the particular path chosen.

12. **Derivatives of lower order.** The proof which we give of the following theorem is only slightly different from that of Montel [1] for functions of a real variable, but it is fundamental in the development and there are some changes which should be pointed out:

THEOREM 12.1. *Let C be an arc or curve of Type W. Let f(z) be continuous* on C and admit a bounded generalized derivative of order $\alpha>0$ on the set C^m, i.e., the set z: z on C, $|z-k|>m>0$, where m is a preassigned positive quantity and k is an arbitrary but fixed point on C. Then f(z) admits a bounded derivative of any order $\alpha'<\alpha$ on C^m.*

CASE I. $0<\alpha'<\alpha<1$. By definition

$$D_z^\alpha f(z) = F'(z), \qquad F(z) = \frac{1}{\Gamma(1-\alpha)}\int_k^z (z-x)^{-\alpha} f(x)dx.$$

Now form

$$D_z^{\alpha'-\alpha} F'(z) = \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F'(x')dx';$$

this expression is the derivative of

$$\begin{aligned} I &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z dz'\int_k^{z'} (z'-x')^{\alpha-\alpha'-1}F'(x')dx' \\ &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z F'(x')dx'\int_{x'}^z (z'-x')^{\alpha-\alpha'-1}dz'. \end{aligned}$$

We can change the order of integration since the curve is of Type W and we have absolute convergence. Hence

$$I = \frac{1}{(\alpha-\alpha')\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'}F'(x')dx'.$$

Integration by parts is permissible and we have

$$\begin{aligned} I &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F(x')dx' \\ &= \frac{1}{\Gamma(\alpha-\alpha')\Gamma(1-\alpha)}\int_k^z (z-x')^{\alpha-\alpha'-1}dx'\int_k^{x'} (x'-x)^{-\alpha}f(x)dx. \end{aligned}$$

Since $f(x)$ is continuous and C is of Type W we may change the order of integration and thus

* For more general functions of a real variable see Tamarkin [1], Theorem 6, p. 227.

$$I = \frac{1}{\Gamma(\alpha - \alpha')\Gamma(1 - \alpha)} \int_k^z f(x)dx \int_x^z (x' - x)^{-\alpha}(z - x')^{\alpha-\alpha'-1}dx'.$$

Now make the transformation $x' - x = t(z - x)$, then

$$I = \frac{1}{\Gamma(\alpha - \alpha')\Gamma(1 - \alpha)} \int_k^z (z - x)^{-\alpha'} f(x)dx \int_0^1 (1 - t)^{\alpha-\alpha'-1} t^{-\alpha} dt.$$

It is well known that

$$(12.10)\quad J = \int_0^1 (1 - t)^{\alpha-\alpha'-1} t^{-\alpha} dt = \frac{\Gamma(\alpha - \alpha')\Gamma(1 - \alpha)}{\Gamma(1 - \alpha')} = B(1 - \alpha, \alpha - \alpha'),$$

where the integral is evaluated along the axis of reals between $t=0$ and $t=1$. If the arc in question does not intersect the segment (0, 1) or intersects it only a finite number of times, we can use the same method as that employed in the preceding paragraph. If there are an infinite number of intersections, we can draw a circle of radius $\frac{1}{2}$ about the point $t=\frac{1}{2}$; it is clear from our previous results that J evaluated over either the upper or lower half of this circle is the same as J evaluated over the segment (0, 1). Also the original arc cannot intersect both the segment (0, 1) and the semi-circle an infinite number of times except in a neighborhood of 0 and of 1 since it is by hypothesis rectifiable. In these neighborhoods we can take the limit since all three arcs are of Type W and the integrals over all three arcs approach 0 as the lengths of the arcs approach 0. Thus we have (12.10) in all cases.

Hence

$$I = \frac{1}{\Gamma(1 - \alpha')} \int_k^z (z - x)^{-\alpha'} f(x)dx,$$

and we know by hypothesis that I admits a bounded derivative

$$\frac{dI}{dz} = D_z^{\alpha'} f(z) = \frac{1}{\Gamma(\alpha - \alpha')} \int_k^z (z - x')^{\alpha-\alpha'-1} F'(x')dx'.$$

Case II. $0 \leqq p - 1 \leqq \alpha' < \alpha < p$. By definition

$$D_z^\alpha f(z) = F^{(p)}(z),^*$$

where

$$F(z) = \frac{1}{\Gamma(p - \alpha)} \int_k^z (z - x)^{p-\alpha-1} f(x)dx.$$

Let

* The expression $f^{(p)}(z)$, where p is a positive integer, denotes the pth derivative of $f(z)$; $f^{(0)}(z) \equiv f(z)$.

$$\Phi(z) = D_z^{\alpha'-\alpha}F^{(p)}(z) = \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F^{(p)}(x')dx',$$

then

$$\begin{aligned}\int_k^z \Phi(z')dz' &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z dz'\int_k^{z'} (z'-x')^{\alpha-\alpha'-1}F^{(p)}(x')dx' \\ &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z F^{(p)}(x')dx'\int_{x'}^z (z'-x')^{\alpha-\alpha'-1}dz' \\ &= \frac{1}{(\alpha-\alpha')\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'}F^{(p)}(x')dx'.\end{aligned}$$

Integration by parts (see above) yields

$$\begin{aligned}\int_k^z \Phi(z')dz' = &-\frac{(z-k)^{\alpha-\alpha'}F^{(p-1)}(k)}{\Gamma(\alpha-\alpha'+1)} \\ &+\frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F^{(p-1)}(x')dx';\end{aligned}$$

and likewise

$$\begin{aligned}\int_k^z dz\int_k^z \Phi(z')dz' = &-\frac{(z-k)^{\alpha-\alpha'+1}F^{(p-1)}(k)}{\Gamma(\alpha-\alpha'+2)}-\frac{(z-k)^{\alpha-\alpha'}F^{(p-2)}(k)}{\Gamma(\alpha-\alpha'+1)} \\ &+\frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F^{(p-2)}(x')dx'.\end{aligned}$$

Finally

$$\begin{aligned}\underbrace{\int_k^z dz\int\cdots\int_k^z}_{p-1}\Phi(z')dz' = &(z-k)^{\alpha-\alpha'}Q_{p-1}(z) \\ &+\frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F(x')dx',\end{aligned}$$

where $Q_{p-1}(z)$ is a polynomial of degree $p-1$ in z. Then

$$\begin{aligned}I &= \frac{1}{\Gamma(\alpha-\alpha')}\int_k^z (z-x')^{\alpha-\alpha'-1}F(x')dx' \\ &= \frac{1}{\Gamma(\alpha-\alpha')\Gamma(p-\alpha)}\int_k^z (z-x')^{\alpha-\alpha'-1}dx'\int_k^{x'} (x'-x)^{p-\alpha-1}f(x)dx \\ &= \frac{1}{\Gamma(\alpha-\alpha')\Gamma(p-\alpha)}\int_k^z f(x)dx\int_x^z (z-x')^{\alpha-\alpha'-1}(x'-x)^{p-\alpha-1}dx'.\end{aligned}$$

By making the same change of variable as above we get

$$I = \frac{1}{\Gamma(p-\alpha')}\int_k^z (z-x)^{p-\alpha-1} f(x)dx,$$

and thus

$$\frac{d^pI}{dz^p} = \Phi(z) - \frac{d^p}{dz^p}\left[(z-k)^{\alpha-\alpha'}Q_{p-1}(z)\right] = D_z^{\alpha'} f(z),$$

and this is bounded in modulus since $|z-k|>m>0$.

CASE III. $\alpha=p$. For this case put

$$\Phi(z) = \frac{1}{\Gamma(p-\alpha)}\int_k^z (z-x)^{p-\alpha'-1} f^{(p)}(x)dx,$$

and

$$I = \frac{1}{\Gamma(p-\alpha')}\int_k^z (z-x)^{p-\alpha'-1} f(x)dx, \quad p-1 \leqq \alpha' < p.$$

Then

$$\frac{d^pI}{dz^p} = \Phi(z) + \frac{d^p}{dz^p}\left[(z-k)^{p-\alpha'}Q_{p-1}(z)\right] = D_z^{\alpha'} f(z).$$

Thus the proof is complete for $\alpha>0$ and $\alpha-\alpha'\leqq 1$. If $\alpha-\alpha'>1$, we can take $\alpha''=p-1$ and by applying the above result a finite number of times obtain the result for an arbitrary $\alpha'<\alpha$.

13. **General properties.** Using methods similar to the ones above we will prove

THEOREM 13.1. *Let C be an arc of Type W. Let $f(z)$ be continuous on C and have a bounded derivative of order $\alpha+\beta$ on C^m. Then we have*

$$D_z^\alpha\{D_z^\beta f(z)\} = D_z^{\alpha+\beta} f(z), \quad \beta < 0, \quad z \text{ on } C^m.$$

For $\alpha=0$ the result is immediate. For $\alpha<0$ we have

$$\begin{aligned} D_z^\alpha\{D_z^\beta f(z)\} &= \frac{1}{\Gamma(-\alpha)\Gamma(-\beta)}\int_k^z (z-x')^{-\alpha-1}dx' \int_k^{x'} (x'-x)^{-\beta-1} f(x)dx \\ &= \frac{1}{\Gamma(-\alpha)\Gamma(-\beta)}\int_k^z f(x)dx \int_x^z (x'-x)^{-\beta-1}(z-x')^{-\alpha-1}dx' \\ &= \frac{1}{\Gamma(-\alpha-\beta)}\int_k^z (z-x)^{-\alpha-\beta-1} f(x)dx = D_z^{\alpha+\beta} f(z). \end{aligned}$$

For $0\leqq p-1\leqq\alpha<p$, we have

$$D_z^\alpha\{D_z^\beta f(z)\} = \frac{d^p}{dz^p} D_z^{\alpha-p}\{D_z^\beta f(z)\} = \frac{d^p}{dz^p} D_z^{\alpha+\beta-p} f(z).$$

If $0 \leqq p-1 \leqq \alpha+\beta < p$, the result follows by definition; if this is not the case, the exponent of $(z-x)$ in the integrand is positive and we can remove the additional integers by the ordinary rules of differentiation.

This theorem is proved by Levy [1] for C a line segment and for $\alpha<0$, $\beta<0$.

Our next result is

THEOREM 13.2. *Let C be an arc of Type W. Let $f(z)$ be continuous on C and have a bounded derivative of order α, $0<\alpha\leqq 1$, on C^m. Also let $f(k)=0$. Then we have*

$$D_z^{-\alpha}\{D_z^\alpha f(z)\} = f(z), \qquad z \text{ on } C^m.$$

For $\alpha=1$ the result is immediate. For $\alpha<1$ we have by definition

$$D_z^\alpha f(z) = \frac{d}{dz} D_z^{\alpha-1} f(z) = \frac{d}{dz}\frac{1}{\Gamma(1-\alpha)}\int_k^z (z-x)^{-\alpha} f(x)dx.$$

Let

$$F(z) = \frac{1}{\Gamma(1-\alpha)}\int_k^z (z-x)^{-\alpha} f(x)dx.$$

By hypothesis $F'(z)$ is bounded. Let

$$\Phi(z) = D_z^{-\alpha}F'(z) = \frac{1}{\Gamma(\alpha)}\int_k^z (z-x')^{\alpha-1}F'(x')dx'.$$

The function $\Phi(z)$ is the derivative of

$$\begin{aligned} I &= \frac{1}{\Gamma(\alpha)}\int_k^z dz\int_k^z (z-x')^{\alpha-1}F'(x')dx' \\ &= \frac{1}{\Gamma(\alpha)}\int_k^z F'(x')dx'\int_{x'}^z (z-x')^{\alpha-1}dz \\ &= \frac{1}{\alpha\Gamma(\alpha)}\int_k^z (z-x')^\alpha F'(x')dx', \end{aligned}$$

and integration by parts yields

$$I = (z-x')^\alpha F(x')\Big|_k^z + \frac{1}{\Gamma(\alpha)}\int_k^z (z-x')^{\alpha-1}F(x')dx'.$$

We have $F(k)=0$ as defined by $\lim_{z\to k}F(z)$. Hence

$$I = \frac{1}{\Gamma(\alpha)} \int_k^z (z - x')^{\alpha-1} F(x') dx'$$

$$= \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_k^z (z - x')^{\alpha-1} dx' \int_k^{x'} (x' - x)^{-\alpha} f(x) dx$$

$$= \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_k^z f(x) dx \int_x^z (z - x')^{\alpha-1}(x' - x)^{-\alpha} dx' = \int_k^z f(x) dx.$$

Thus

$$D_z^{-\alpha} F'(z) = D_z^{-\alpha}\{D_z^\alpha f(z)\} = \Phi(z) = \frac{dI}{dz} = \frac{d}{dz}\int_k^z f(x)dx = f(z).$$

For $\alpha > 1$ additional conditions on the original function $f(z)$ are involved. The same method can be used to investigate this case. If $f(k) = f'(k) = \cdots = f^{(n)}(k) = 0$, and if $f(z)$ has bounded ordinary derivatives of sufficiently high order Theorem 13.2 can be extended to values of $\alpha \leqq n$. The case where $0 < \alpha \leqq 1$ is the most interesting, consequently the details of the extension will not be included here.

As an immediate consequence of the definition we have

THEOREM 13.3. *Let the functions $f_1(z)$ and $f_2(z)$ be continuous and have bounded derivatives of order $\alpha > 0$ on C^m, where C is an arc of Type W. Then we have*

$$D_z^\alpha\{f_1(z) + f_2(z)\} = D_z^\alpha f_1(z) + D_z^\alpha f_2(z).$$

We have also

THEOREM 13.4. *Let C be an arc of Type W. Let $f(z)$ be continuous on C and have a bounded derivative of order α, $0 \leqq p-1 \leqq \alpha < p$, on C^m. Then $f^{(p-1)}(z)$ has a bounded derivative of order $\alpha - p + 1$ on C^m.*

By Theorem 12.1 we know that $f^{(p-1)}(z)$ exists and is bounded and by definition

$$D_z^{\alpha-p+1} f^{(p-1)}(z) = \frac{d}{dz} D_z^{\alpha-p} f^{(p-1)}(z)$$

$$= \frac{d}{dz}\left\{\frac{1}{\Gamma(p-\alpha)}\int_k^z (z - x)^{p-\alpha-1} f^{(p-1)}(x) dx\right\}.$$

We have to show that

$$\Phi(z) = \frac{1}{\Gamma(p-\alpha)}\int_k^z (z - x)^{p-\alpha-1} f^{(p-1)}(x) dx$$

has a bounded derivative, with the hypothesis that $\int_k^z (z-x)^{p-\alpha-1} f(x)dx$ has a bounded pth derivative. If we can prove that

$$\underbrace{\int_k^z dz \int \cdots \int_k^z}_{p-1} \Phi(z)dz$$

has a bounded pth derivative, the theorem is established. We have

$$\frac{1}{\Gamma(p-\alpha)} \int_k^z dz \int_k^z (z-x)^{p-\alpha-1} f^{(p-1)}(x)dx$$

$$= \frac{1}{\Gamma(p-\alpha)} \int_k^z f^{(p-1)}(x)dx \int_x^z (z-x)^{p-\alpha-1} dz$$

$$= \frac{1}{(p-\alpha)\Gamma(p-\alpha)} \int_k^z (z-x)^{p-\alpha} f^{(p-1)}(x)dx,$$

$$= -\frac{(z-k)^{p-\alpha} f^{(p-2)}(k)}{\Gamma(p-\alpha+1)} + \frac{1}{\Gamma(p-\alpha)} \int_k^z (z-x)^{p-\alpha-1} f^{(p-2)}(x)dx,$$

integrating by parts. Likewise

$$\int_k^z dz \int_k^z \Phi(z)dz = -\frac{(z-k)^{p-\alpha+1} f^{(p-2)}(k)}{\Gamma(p-\alpha+2)} - \frac{(z-k)^{p-\alpha} f^{(p-3)}(k)}{\Gamma(p-\alpha+1)}$$

$$+ \frac{1}{\Gamma(p-\alpha)} \int_k^z (z-x)^{p-\alpha-1} f^{(p-3)}(x)dx.$$

Finally

$$\underbrace{\int_k^z dz \int \cdots \int_k^z}_{p-1} \Phi(z)dz = (z-k)^{p-\alpha} Q_{p-2}(z)$$

$$+ \frac{1}{\Gamma(p-\alpha)} \int_k^z (z-x)^{p-\alpha-1} f(x)dx.$$

The function $Q_{p-2}(z)$ is a polynomial of degree $p-2$ in z. Since the expression on the right has a bounded pth derivative for $|z-k| > m > 0$, the expression on the left does also and the proof is complete.

A related theorem is the following:

THEOREM 13.5. *Let C be an arc of Type W. Let $f^{(p-1)}(z)$ be continuous on C and have a bounded derivative of order $\alpha-p+1$, $0 \leq p-1 \leq \alpha < p$, on C^m. Then $f(z)$ has a bounded derivative of order α on C^m.*

We know that

$$\frac{d}{dz}\int_k^z (z-x)^{p-\alpha-1}f^{(p-1)}(x)dx$$

is bounded and we have to prove that

$$\frac{d^p}{dz^p}\int_k^z (z-x)^{p-\alpha-1}f(x)dx$$

is bounded. Integrating by parts we have

$$\Phi(z) = \int_k^z (z-x)^{p-\alpha-1}f(x)dx = \frac{(z-k)^{p-\alpha}f(k)}{p-\alpha} + \int_k^z (z-x)^{p-\alpha}f'(x)dx,$$

and integrating the second integral on the right by parts yields

$$\Phi(z) = \frac{(z-k)^{p-\alpha}f(k)}{p-\alpha} + \frac{(z-k)^{p-\alpha+1}f'(k)}{p-\alpha+1} + \int_k^z (z-x)^{p-\alpha+1}f^{(2)}(x)dx.$$

Continuing this process we arrive at

$$\Phi(z) = \frac{(z-k)^{p-\alpha}f(k)}{p-\alpha} + \frac{(z-k)^{p-\alpha+1}f'(k)}{p-\alpha+1} + \cdots$$
$$+ \frac{(z-k)^{2p-\alpha-2}f^{(p-2)}(k)}{2p-\alpha-2} + \int_k^z (z-x)^{2p-\alpha-2}f^{(p-1)}(x)dx.$$

We know that each of the terms on the right exclusive of the integral has a bounded pth derivative on C^m and thus there remains only to show that this integral has a bounded pth derivative. We can differentiate under the integral sign $p-1$ times and the result is

$$\int_k^z (z-x)^{p-\alpha-1}f^{(p-1)}(x)dx.$$

By hypothesis this has a bounded first derivative and the result is established. There are similar theorems for $f^{(p-2)}(z)$, etc.

The proof of the following theorem is similar in method to that used in showing that the *Beta function* [see (12.10)] can be evaluated along curves of Type W (see also §11):

THEOREM 13.6. *Let $f(z)$ be analytic in a region bounded by a curve C of Type W and continuous in the corresponding closed region except for a finite number of algebraic singularities of orders less than* 1. *Then for a particular branch of the function we have $\int_C f(z)dz=0$.*

14. **Invariance of order under conformal transformation.** Let C be an analytic Jordan curve in the z-plane and let k be a fixed point on C. Let

$w=\phi(z)$ map the closed region bounded by C conformally on the closed unit circle in the w-plane. It is well known that $\phi(z)$ is analytic on C, and that $z=\psi(w)$, the inverse of $w=\phi(z)$, is analytic on $|w|=1$. *Let $f(z)$ be continuous on C and have a bounded derivative of order $\alpha>0$ on C^m. We will show that $f(\psi(w))=F(w)$ has a bounded derivative of the same order α on the arc of the unit circle, $|w|=1$, corresponding to C^m under the conformal map, the path of integration being the image in the w-plane of the path of integration in the z-plane.*

CASE I. $0<\alpha<1$. By definition

$$D_z^\alpha f(z) = \frac{d}{dz}\frac{1}{\Gamma(1-\alpha)}\int_k^z (z-x)^{-\alpha}f(x)dx,$$

and by hypothesis

$$g(z) = \int_k^z (z-x)^{-\alpha}f(x)dx$$

has a bounded first derivative with respect to z. Then

$$\begin{aligned} g(\psi(w)) &= \int_l^w [\psi(w)-\psi(\omega)]^{-\alpha}F(\omega)\psi'(\omega)d\omega \\ &= \int_l^w (w-\omega)^{-\alpha}\left[\frac{\psi(w)-\psi(\omega)}{w-\omega}\right]^{-\alpha}F(\omega)\psi'(\omega)d\omega \end{aligned}$$

has a bounded first derivative with respect to w, where k on C is mapped into l on $|w|=1$. The function

$$P(w,\omega) = \left[\frac{\psi(w)-\psi(\omega)}{w-\omega}\right]^{-\alpha}\psi'(\omega)$$

is an analytic function of w and ω and is different from 0. We may write it in the form

$$\begin{aligned} P(w,\omega) &= P(w,w) + \frac{P(w,\omega)-P(w,w)}{w-\omega}(w-\omega) \\ &= P(w,w) + Q(w,\omega)(w-\omega), \end{aligned}$$

where $P(w,w)\neq 0$ and is defined as the limit of $P(w,\omega)$ as ω approaches w. Thus we have

$$\begin{aligned} g(\psi(w)) = G(w) &= \int_l^w (w-\omega)^{-\alpha}[P(w,w)+Q(w,\omega)(w-\omega)]F(\omega)d\omega \\ &= P(w,w)\int_l^w (w-\omega)^{-\alpha}F(\omega)d\omega + \int_l^w (w-\omega)^{-\alpha+1}Q(w,\omega)F(\omega)d\omega. \end{aligned}$$

By hypothesis $G(w)$ has a bounded derivative with respect to w. Now set

$$A(w) = \int_l^w (w - \omega)^{-\alpha+1} Q(w, \omega) F(\omega) d\omega.$$

By Leibnitz' rule we have

$$\frac{dA(w)}{dw} = \int_l^w (w - \omega)^{-\alpha+1} Q_w(w, \omega) F(\omega) d\omega$$

$$+ \int_l^w (-\alpha + 1)(w - \omega)^{-\alpha} Q(w, \omega) F(\omega) d\omega.$$

The integral on the right converges absolutely and hence the derivative exists. It follows that $P(w, w)\int_l^w (w-\omega)^{-\alpha} F(\omega) d\omega$ has a bounded derivative with respect to w and, since $P(w, w) \neq 0$ and has a derivative, we have shown that for $0 < \alpha < 1$ the function $F(w)$ has a bounded derivative of order α.

CASE II. $0 \leqq p-1 \leqq \alpha < p$. If $\alpha = p-1$ the result is immediate. If $\alpha > p-1$ we know by Theorem 12.1 that $f^{(p-1)}(z)$ is bounded and it follows that $F^{(p-1)}(w)$ is bounded. In fact we know that $f^{(p-1)}(z)$ has a bounded derivative of order $\alpha - p + 1$ by Theorem 13.4. The expression for $F^{(p-1)}(w)$ is a polynomial in $f^{(1)}, f^{(2)}, \cdots, f^{(p-1)}, \psi^{(1)}, \psi^{(2)}, \cdots, \psi^{(p-1)}$, and we can treat each separate integral as in Case I and the proof is complete.

15. **Extension of the Cauchy integral formula.** We obtain a new expression for the generalized derivative from the following

THEOREM 15.1. *Let C be an arc of Type W, k a fixed point on C, and z an arbitrary point on C, $z \neq k$. Let $f(z)$ be analytic on and within a rectifiable Jordan curve γ which passes through k and contains in its interior that portion of C between k and z. Then we have*

$$D_z^\alpha f(z) = \frac{\Gamma(\alpha + 1)}{2\pi i} \int_\gamma \frac{f(t)dt}{(t - z)^{\alpha+1}}, \qquad \alpha > 0, \tag{15.10}$$

where the branch of $(t-z)^{\alpha+1}$ is the one chosen in the original definition with t varying along C from k to z.

This is a generalization of Cauchy's integral formula for derivatives and the proof follows that of Montel [1] for C a segment of the axis of reals. By virtue of (11.14) for C of Type W we may replace it by C_1 of Type W, and thus it is only necessary that γ pass through k and contain in its interior the point z and an arc of Type W joining k to z, so long as $f(z)$ is analytic on and within γ, in fact it is sufficient for $f(z)$ to be analytic interior to γ and continuous in the closed limited region bounded by γ. For future reference we will state the following more general theorem:

THEOREM 15.2. *Let $f(z)$ be analytic in the interior of a rectifiable Jordan curve γ which passes through k and contains in its interior the point z and an arc C of Type W joining k to z, and let $f(z)$ be continuous in the closed limited region bounded by γ. Then we have*

$$D_z^\alpha f(z) = \frac{\Gamma(\alpha+1)}{2\pi i}\int_\gamma \frac{f(t)dt}{(t-z)^{\alpha+1}}, \qquad \alpha > 0, \tag{15.11}$$

where the branch of $(t-z)^{\alpha+1}$ is determined on C.

16. **A theorem on convergence.** The following theorem has important application later:

THEOREM 16.1. *Let C be an arc or curve of Type W. Let $f_1(z), f_2(z), \cdots$ be a sequence of uniformly bounded functions which converges uniformly to $f(z)$ on C. Let $f_1^\alpha(z)$, $f_2^\alpha(z), \cdots$, $\alpha>0$, converge uniformly on C^m. Then the sequence $f_1^\alpha(z), f_2^\alpha(z), \cdots$ has for a limit $f^\alpha(z)$ on C^m.*

Montel [1] proves this theorem for C a segment of the axis of reals and the modifications necessary to establish this more general result are obvious, hence the proof will not be included here.

17. **Lipschitz conditions and generalized derivatives.** We begin with the following theorem:*

THEOREM 17.1. *Let C be an arc or curve of Type W. Let $f(z)$ satisfy a Lipschitz condition† of order α, $1 \geqq \alpha > 0$, on C. Then $f(z)$ has a bounded derivative of every order $\beta < \alpha$ on C^m.*

For $\beta \leqq 0$ the result is obvious. For $\beta > 0$ we have to show that

$$\int_k^z (z-x)^{-\beta} f(x)dx$$

has a bounded derivative. Consider

$$I(z) = \int_k^{z-\epsilon} (z-x)^{-\beta} f(x)dx.$$

$$\frac{dI}{dz} = \epsilon^{-\beta} f(z-\epsilon) - \beta\int_k^{z-\epsilon}(z-x)^{-\beta-1}f(x)dx$$

$$= \epsilon^{-\beta}[f(z-\epsilon) - f(z)] + \beta\int_k^{z-\epsilon}[f(z) - f(x)](z-x)^{-\beta-1}dx.$$

* For the case where C is a straight line the proof is due to Hardy and Littlewood [1], pp. 590–591.

† The function $f(z)$ satisfies a Lipschitz condition of order α on a set E if for z_1 and z_2 arbitrary points on E we have $|f(z_1)-f(z_2)| \leqq M|z_1-z_2|^\alpha$, where M is a constant independent of z_1 and z_2.

The first term on the right approaches 0 as ϵ approaches 0 by the hypothesis of a Lipschitz condition of order α on the function $f(z)$. Furthermore

$$\left|\int_k^{z-\epsilon}[f(z)-f(x)](z-x)^{-\beta-1}dx\right| \leqq \int_k^{z-\epsilon}|f(z)-f(x)|\,|z-x|^{-\beta-1}|dx|$$
$$\leqq \int_k^{z-\epsilon} M|z-x|^{\alpha}|z-x|^{-\beta-1}|dx|$$
$$= M\int_k^{z-\epsilon}|z-x|^{\alpha-\beta-1}|dx|.$$

The last integral converges and the proof is complete.

Now we will consider the case where $f^{(p-1)}(z)$ satisfies a Lipschitz condition of order $\alpha-p+1$, $0\leqq p-1\leqq\alpha<p$, on C. From above we know that $f^{(p-1)}(z)$ has a bounded derivative of every order $\beta-p+1<\alpha-p+1$, and hence by Theorem 13.5 the function $f(z)$ has a bounded derivative of every order $\beta<\alpha$ on C^m. Consequently we can state the following more general theorem:

THEOREM 17.2. *Let C be a curve of Type W. Let $f^{(p-1)}(z)$ satisfy a Lipschitz condition of order $\alpha-p+1$, $0\leqq p-1\leqq\alpha<p$, on C. Then $f^{(p-1)}(z)$ has a bounded derivative of every order $\beta<\alpha$ on C^m.*

In considering the converse we restrict ourselves to analytic Jordan curves.

THEOREM 17.3. *Let C be an analytic Jordan curve. Let $f(z)$ have a bounded derivative of order $\alpha>0$, $0\leqq p-1\leqq\alpha<p$, on C^m. Then $f^{(p-1)}(z)$ satisfies a Lipschitz condition of order $\alpha-p+1$ on C.*

Since Lipschitz conditions and generalized derivatives are preserved in their orders under conformal transformation of analytic arcs and curves (see §14), for $0<\alpha\leqq 1$ the theorem is an immediate consequence of the corresponding result of Weyl [1] for C a segment of the axis of reals. For the general α by Theorem 13.4 we know that $f(z)$ has a bounded derivative of order $\alpha-p+1$, and hence $f^{(p-1)}(z)$ satisfies a Lipschitz condition of order $\alpha-p+1$.

It should be noted here that we must use two different values of k to obtain an overlapping Lipschitz condition. This overlap gives us a Lipschitz condition which is uniform on C.

Suppose that C is a Jordan curve composed of a finite number of analytic Jordan arcs meeting in corners of openings $\mu_k\pi$, $k=1, 2, \cdots, p$, $0<\mu_k<2$. By the above result we know that if $f(z)$ has a bounded derivative of order $\alpha>0$ on C^m, then $f^{(p-1)}(z)$ satisfies a uniform Lipschitz condition of order $\alpha-p+1$ on each analytic arc of C. The following lemma serves to connect Lipschitz conditions on two adjacent arcs:

LEMMA 17.4. *Let C be a Jordan arc composed of two analytic arcs meeting in a corner z_0 of opening $\mu\pi$, $0<\mu<2$. Let the function $f(z)$ satisfy a uniform Lipschitz condition of order α, $0<\alpha\leqq 1$, on each arc separately. Then $f(z)$ satisfies a uniform Lipschitz condition of order α on the entire arc.*

The lemma is an immediate consequence of the convexity of the function $y=x^\alpha$, $x>0$.

CHAPTER IV

GENERALIZATIONS OF BERNSTEIN'S AND MARKOFF'S THEOREMS

18. **Bernstein's and Markoff's theorems.** The following theorem is well known:*

THEOREM 18.1. *Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)|\leqq M$ in the unit circle, $|z|\leqq 1$. Then $|P_n'(z)|\leqq Mn$ in $|z|\leqq 1$. This bound is attained only by the polynomial αz^n, $|\alpha|=1$.*

Markoff [1]† has proved

THEOREM 18.2. *Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)|\leqq M$ in the interval $-1\leqq z\leqq +1$. Then $|P_n'(z)|\leqq Mn^2$, $-1\leqq z\leqq +1$. This bound is attained only by the polynomial α cos n arc cos z, $|\alpha|=1$.*

A result of Bernstein [2, p. 38] on the modulus of the derivative of a polynomial is

THEOREM 18.3. *Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)|\leqq M$ in the interval $-1\leqq z\leqq +1$. Then $|P_n'(z)|\leqq Mn/(1-z^2)^{1/2}$.*

Montel [1] has extended this to generalized derivatives.

THEOREM 18.4. *Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)|\leqq M$, $-1\leqq z\leqq +1$. Then for any point in $(-z_0, +z_0)$, $0<z_0<1$, we have*

$$(18.10)\qquad |P_n^\alpha(z)|\leqq \frac{M(\alpha)n^\alpha}{(1-z_0^2)^{\alpha/2}},\qquad \alpha>0,\qquad k=+1.$$

We are concerned with finding evaluations for $|P_n^\alpha(z)|$ on more general point sets. It should be noted that inequality (18.10) is not valid in the entire interval but only on a subset of the given interval. This is characteristic of the results on generalized derivatives. In fact in this chapter our inequalities hold in general only for $|z-k|>m>0$, and in the case of a curve C we shall

---

* See e.g., Bernstein [2], pp. 44–46.

† Markoff considers only polynomials with real coefficients; Riesz [1] treats the general case. See e.g., Bernstein [2], p. 38.

denote the set: z on C, $|z-k|>m>0$, by C^m, as above, and in the case of a region R the set: z in $\bar{R}$, $|z-k|>m>0$, by $\bar{R}^m$.

19. **Curves of Type S.** Let C be a curve of Type S in the z-plane and let $|P_n(z)|\leqq M$ on C. We shall consider $P_n^\alpha(z)$, $\alpha>0$, on C, where k is an arbitrary point on C. By Theorem 15.2 we know that

$$P_n^\alpha(z)=\frac{\Gamma(\alpha+1)}{2\pi i}\int_\gamma\frac{P_n(t)dt}{(t-z)^{\alpha+1}},\qquad \alpha>0, \tag{19.10}$$

where γ is a rectifiable Jordan curve passing through k and containing z in its interior. Let z_0 be an arbitrary point on C^m and draw a circle δ of radius $r<m$ about the point z_0, the value of r to be determined later. In traversing the curve C in a preassigned direction from k to z let the point A be the first intersection of δ with C. We will choose as the curve γ the path from k to A, from A around δ back to A, and thence back to k along C. It is true that this is not a Jordan curve but reference to the proof of Montel [1] of Theorem 15.1 will make it clear that this path of integration satisfies the conditions of the theorem. Then we have

$$P_n^\alpha(z_0)=\frac{\Gamma(\alpha+1)}{2\pi i}\int_k^A\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}+\frac{\Gamma(\alpha+1)}{2\pi i}\int_\delta\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}$$
$$+\frac{\Gamma(\alpha+1)e^{-2\pi i(\alpha+1)}}{2\pi i}\int_k^A\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}.$$

Thus

$$|P_n^\alpha(z_0)|\leqq\frac{\Gamma(\alpha+1)}{2\pi}\left\{2\left|\int_k^A\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}\right|+\left|\int_\delta\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}\right|\right\}.$$

Let

$$I'\equiv\int_\delta\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}\quad\text{and}\quad I''\equiv\int_k^A\frac{P_n(t)dt}{(t-z_0)^{\alpha+1}}.$$

We need a well known result,*

LEMMA 19.1. *Let E be a closed limited point set whose complement is connected and regular. If the polynomial $P_n(z)$ of degree n in z satisfies the inequality $|P_n(z)|\leqq M$ for z on E, then we have*

$$|P_n(z)|\leqq M\rho^n,\quad \rho>1,\quad z\text{ on or within }C_\rho. \tag{19.11}$$

We know by Theorem 3.1 that $d(C,C_\rho)\geqq M_1(\rho-1)$ and hence if we choose $r=M_1(\rho-1)$ it follows from the lemma that $|P_n(z)|\leqq M\rho^n$ for z on and within the circle δ. Then we have

* For various types of point sets this lemma is due to various authors: see Bernstein [1], Szegö [1], Faber [4], Walsh [1]. We state here the form due to Walsh.

$$|I'| \leqq \int_s \frac{|P_n(t)||dt|}{|t-z_0|^{\alpha+1}} \leqq \frac{M\rho^n 2\pi M_1(\rho-1)}{M_1^{\alpha+1}(\rho-1)^{\alpha+1}} = \frac{2\pi M}{M_1^\alpha} \frac{\rho^n}{(\rho-1)^\alpha}.$$

The function $y=\rho^n/(\rho-1)^\alpha$, $\rho>1$, has a minimum for $\rho=\rho_n\equiv n/(n-\alpha)$. For n sufficiently large $M_1(\rho_n-1)=M_1[\alpha/(n-\alpha)]$ will be positive and less than m, in fact our evaluation will hold for all n satisfying the inequality

$$(19.12) \qquad M_1\left(\frac{\alpha}{n-\alpha}\right) < m, \quad (n-\alpha) > \frac{M_1\alpha}{m}, \quad n > \frac{M_1\alpha}{m} + \alpha.$$

Now consider $\rho_n{}^n/[(\rho_n-1)^\alpha n^\alpha]$. It is easy to see that this function approaches $e^{2\alpha}\alpha^{-\alpha}$ as n approaches ∞ and hence we have

$$(19.13) \qquad |I'| \leqq MK_1(\alpha, C)n^\alpha.$$

Next

$$|I''| \leqq M\int_k^A \frac{|dt|}{|t-z_0|^{\alpha+1}}.$$

By §10 the curve C is of Type W', and since A is on the circumference of a circle whose center is z_0 and whose radius is $M_1[\alpha/(n-\alpha)]$, we have

$$(19.14) \qquad |I''| \leqq MM_1\left\{N_1\left(\frac{\alpha}{n-\alpha}\right)^{-\alpha} + N_2\right\} \leqq MK_2(\alpha, C)n^\alpha,$$

which is valid for n satisfying the inequality (19.12).

Thus we have established the following theorem:

THEOREM 19.2. *Let C be a curve of Type S. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ on C. Then for n sufficiently large we have*

$$(19.15) \qquad |P_n^\alpha(z)| \leqq MK(\alpha, C)n^\alpha, \quad \alpha > 0, \quad z \text{ on } C^m,$$

where $K(\alpha, C)$ is a constant depending only on α and C.

20. **Curves of Type t.** The method of the preceding paragraph is valid in the case where C is a curve of Type t with only slight modification. In fact the only essential change is the substitution of αt for α, since by Theorem 4.3 we have here $d(C, C_\rho) \geqq M_1(\rho-1)^t$. The result can be stated as follows:

THEOREM 20.1. *Let C be a curve of Type t. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ on C. Then for n sufficiently large we have*

$$(20.10) \qquad |P_n^\alpha(z)| \leqq MK(\alpha, C)n^{\alpha t}, \quad \alpha > 0, \quad 1 \leqq t < 2, \quad z \text{ on } C^m,$$

where $K(\alpha, C)$ is a constant depending only on α and C.

21. **Regions of Type W'.** In this case we have $d(C, C_\rho) \geqq M_1(\rho-1)^2$ by Theorem 6.3, and hence the t in Theorem 20.1 must be replaced by 2. The

definition of a region of Type W' (§10) then leads directly by an application of the above methods to

THEOREM 21.1. *Let R be a region of Type W' with boundary C. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ in $\overline{R}$. Then for n sufficiently large we have*

$$(21.10) \qquad |P_n^\alpha(z)| \leqq MK(\alpha, R)n^{2\alpha}, \quad \alpha > 0, \quad z \text{ on } C^m,$$

where $K(\alpha, R)$ is a constant depending only on α and R.

Here k is a point on C and, as in §§19 and 20, $P_n^\alpha(z)$ is a continuous function of z on C^m. There is no confusion as to branches since both k and z are boundary points.

22. **Another choice of k.** In the case where $f(z)$ is analytic in a limited simply connected region R with boundary C, and continuous in the corresponding closed region, we may choose k not on C but in R. Then we can consider all z in $\overline{R}$ such that $|z-k| > m > 0$, i.e., in $\overline{R}^m$. Of course the value of the derivative depends upon the point k chosen.

The original choice of k on C is more general in that the function need be defined only on the set C. However, in the application of these theorems to approximation by polynomials in closed regions the functions are assumed analytic in the regions and continuous in the corresponding closed regions. Under these conditions it is simpler to choose k inside and then we know that the above evaluations hold not only on the boundary but everywhere in the closed region except for the points z such that $|z-k| \leqq m$.

Thus we have the following results:

THEOREM 22.1. *Let R be the limited region bounded by a curve C of Type S. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ in $\overline{R}$. Then for n sufficiently large we have*

$$|P_n^\alpha(z)| \leqq MK(\alpha, R)n^{\alpha}, \quad \alpha > 0, \quad z \text{ in } \overline{R}^m,$$

where $K(\alpha, R)$ depends only on α and R.

THEOREM 22.2. *Let R be the limited region bounded by a curve C of Type t. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ in $\overline{R}$. Then for n sufficiently large we have*

$$|P_n^\alpha(z)| \leqq MK(\alpha, R)n^{\alpha t}, \quad \alpha > 0, \quad 1 \leqq t < 2, \quad z \text{ in } \overline{R}^m,$$

where $K(\alpha, R)$ depends only on α and R.

THEOREM 22.3. *Let R be a region of Type W' with boundary C. Let $P_n(z)$ be a polynomial of degree n in z and let $|P_n(z)| \leqq M$ in $\overline{R}$. Then for n sufficiently large we have*

$$|P_n^\alpha(z)| \leqq MK(\alpha, R)n^{2\alpha}, \quad \alpha > 0, \quad z \text{ in } \overline{R}^m,$$

where $K(\alpha, R)$ depends only on α and R.

If we choose k in R and m sufficiently small the boundary C will lie in $\overline{R}^m$.

23. **Ordinary derivatives.** If α is a positive integer we have to consider only the part I' of the expression for $P_n^\alpha(z)$ in §19. Thus we avoid the question of curves of Type W'. In this connection Szegö [2] has considered the case where C is a Jordan curve composed of a finite number of analytic arcs. He shows that if z_0 is a corner whose exterior opening is $\mu\pi$, $0<\mu\leqq 2$, and if $|P_n(z)|\leqq M$ on C, then $MM_1n^\mu \leqq |P_n'(z_0)| \leqq MM_2n^\mu$, where M_1 and M_2 are constants independent of n (see Theorem 22.2).

Jackson [2] has shown that if the boundary C of a limited simply connected region R is such that through every boundary point P of C a circle of radius $h>0$, h independent of P, can be drawn whose interior lies interior to the region, then $|P_n(z)|\leqq M$, z in $\overline{R}$, implies $|P_n'(z)|\leqq MM_2n$, z in $\overline{R}$, where M_2 is a constant independent of M, n, and z; this result follows from Theorem 18.1. Also Jackson [3] has shown that for a limited simply connected region R such that from every boundary point P a straight line of length $h>0$, h independent of P, can be drawn which lies entirely in R, the inequality $|P_n(z)|\leqq M$, z in $\overline{R}$, implies $|P_n'(z)|\leqq MM_2n^2$, z in $\overline{R}$; it is clear that this is a region of Type W' and hence the result is a corollary of Theorem 21.1. In this connection the author [3] has shown that if $|P_n(z)|\leqq M$ on a set E with no isolated points and whose complement has finite connectivity, then $|P_n'(z)|\leqq MK(E)n^2$, z on E, where the constant $K(E)$ depends only on the set E.

Chapter V

Approximation by polynomials—Problem α

24. **Analytic Jordan curves.** We will prove the following theorem:

Theorem 24.1. *Let R be a limited region bounded by an analytic Jordan curve C. Let $f(z)$ be defined in $\overline{R}$. If for every n there exists a polynomial $P_n(z)$ of degree n in z such that*

$$|f(z) - P_n(z)| < \frac{M}{n^\alpha}, \quad \alpha > 0, \quad z \text{ in } \overline{R}, \tag{24.10}$$

M a constant independent of n and z, then $f(z)$ is analytic in R, continuous in $\overline{R}$, and has a bounded derivative of every order $\alpha'<\alpha$ on C.

The analyticity and continuity are immediate consequences of (24.10). For the derivative we have to apply our previous results. Hereafter we will choose an interior point of the region as the point k and m so small that every

point of C belongs to $\bar{R}^m$. Using Theorems 22.1 and 16.1 we can apply a procedure identical to that of Montel [1, pp. 175–176] to prove the theorem.

We have in the other direction

THEOREM 24.2. *Let R be a limited region bounded by an analytic Jordan curve C. Let $f(z)$ be analytic in R, continuous in $\bar{R}$, and have a bounded derivative of order $\alpha>0$ on C. Then there exist polynomials of respective degrees n, $n=1, 2, \cdots$, such that*

$$|f(z) - P_n(z)| \leqq \frac{M}{n^\alpha}, \qquad z \text{ in } \bar{R}, \tag{24.11}$$

where M is a constant independent of n and z.

Let $0 \leqq p-1 \leqq \alpha < p$, then by Theorem 17.3 we know that $f^{(p-1)}(z)$ satisfies a uniform Lipschitz condition of order $\alpha-p+1$ on C. Thus we have only to apply the following theorem of Curtiss [1]:

THEOREM 24.3. *Let E, with boundary C, be a closed limited point set consisting of a finite number of mutually exterior closed Jordan regions, $R_1, R_2, \cdots, R_s$, each bounded by an analytic Jordan curve. Let $f(z)$ be a function analytic throughout the interior of C, continuous on E, and having a $(p-1)$st derivative which satisfies a Lipschitz condition of order $\alpha-p+1$, $0<\alpha-p+1\leqq 1$, on the boundary of E. Then there exist polynomials $P_n(z)$ of respective degrees n, $n=1, 2, \cdots$, such that*

$$|f(z) - P_n(z)| \leqq \frac{M}{n^\alpha}, \qquad z \text{ on } E,$$

*where M is a constant independent of n and z.**

It should be observed here that Theorem 24.1 taken with Theorem 24.3 leads to a new proof of Theorem 17.1 where C is an analytic Jordan curve.

Also Theorem 17.3 leads to the following result:

THEOREM 24.4. *Under the same hypothesis as in Theorem 24.1, let $0\leqq p-1\leqq\alpha<p$. Then $f(z)$ is analytic in R, continuous in $\bar{R}$, and $f^{(p-1)}(z)$ satisfies a Lipschitz condition of every order $\alpha'-p+1$, $\alpha'<\alpha$, on C.*

25. **Curves of Type** S. In the proof of Theorem 24.1 we apply Theorem 22.1 which does not require C to be analytic but merely of Type S and thus we have the following more general theorem:

THEOREM 25.1. *Let R be a limited region bounded by a curve C of Type S.*

---

* This theorem is a sharpening of a result on degree of convergence of Faber's polynomials established by the author [1].

Let $f(z)$ be defined in $\overline{R}$. If for every n, $n=1, 2, \cdots$, there exists a polynomial $P_n(z)$ of degree n in z such that

$$(25.10) \qquad |f(z) - P_n(z)| \leqq \frac{M}{n^\alpha}, \quad \alpha > 0, \quad z \text{ in } \overline{R},$$

M a constant independent of n and z, then $f(z)$ is analytic in R, continuous in $\overline{R}$, and has a bounded derivative of every order $\alpha' < \alpha$ on C.

Only in the case where C is analytic has the converse theorem been proved.

26. **Curves with corners.** Here we consider first curves of Type t, and using Theorem 22.2 we can prove precisely as above

THEOREM 26.1. *Let R be a limited region bounded by a curve C of Type t. Let $f(z)$ be defined in $\overline{R}$. If for every n, $n=1, 2, \cdots$, there exists a polynomial $P_n(z)$ of degree n in z such that*

$$(26.10) \qquad |f(z) - P_n(z)| \leqq \frac{M}{n^{\alpha t}}, \quad \alpha > 0, \quad 1 \leqq t < 2, \quad z \text{ in } \overline{R},$$

M a constant independent of n and z, then $f(z)$ is analytic in R, continuous in $\overline{R}$, and has a bounded derivative of every order $\alpha' < \alpha$ on C.

Combining the above result with Theorems 17.3 and 17.4 we can substitute a Lipschitz condition for the generalized derivative.

THEOREM 26.2. *Let R be a limited region bounded by a Jordan curve C composed of a finite number of analytic arcs meeting in corners of exterior openings $\mu_k\pi$, $2 > \mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_s > 0$, and let $t = \mu_1$ if $\mu_1 \geqq 1$, and $t=1$ if $\mu_1 < 1$. Let $f(z)$ be defined in $\overline{R}$. If for every n, $n=1, 2, \cdots$, there exists a polynomial $P_n(z)$ of degree n in z such that*

$$|f(z) - P_n(z)| \leqq \frac{M}{n^{\alpha t}}, \quad 0 \leqq p - 1 \leqq \alpha < p, \quad 2 > t \geqq 1, \quad z \text{ in } \overline{R},$$

M a constant independent of n and z, then $f(z)$ is analytic in R, continuous in $\overline{R}$, and $f^{(p-1)}(z)$ satisfies a Lipschitz condition of every order $\alpha' - p + 1$, $\alpha' < \alpha$, on C.

27. **Regions of Type W'.** Here the result corresponding to Theorem 26.1 is

THEOREM 27.1. *Let R be a region of Type W' with boundary C and let $f(z)$ be defined in $\overline{R}$. If for every n, $n=1, 2, \cdots$, there exists a polynomial $P_n(z)$ of degree n in z such that*

$$(27.10) \qquad |f(z) - P_n(z)| \leqq \frac{M}{n^{2\alpha}}, \quad \alpha > 0, \quad z \text{ in } \overline{R},$$

M a constant independent of n and z, then $f(z)$ is analytic in R, continuous in $\overline{R}$, and has a bounded derivative of every order $\alpha' < \alpha$ on C.

28. **Arbitrary closed sets.** Theorem 27.1 can be proved for ordinary derivatives on much more general sets than those of §27. (See Sewell [3].) We do not require the set to be connected; it is sufficient for the set to be closed and limited and have a complement which is connected and regular.

Chapter VI

Approximation by polynomials—Problem β

29. **Uniformly bounded functions.** We begin with

Theorem 29.1. *Let $f(z)$ be analytic and uniformly bounded in $|z| \leqq \rho > 1$. Then we have*

$$|f(z) - T_n(z)| \leqq \frac{M}{\rho^n}, \quad |z| \leqq 1, \tag{29.10}$$

where M is a constant independent of n and z, and $T_n(z)$ is the sum of the first n terms of the Taylor development of $f(z)$ about $z=0$.

We know that

$$f(z) - T_n(z) = \frac{z^{n+1}}{2\pi i}\int_{|z|=\rho-\epsilon} \frac{f(t)dt}{t^{n+1}(t-z)}, \quad \rho - 1 > \epsilon > 0, \quad |z| \leqq 1;$$

hence

$$|f(z) - T_n(z)| \leqq \frac{1}{2\pi}\,\frac{M_1 2\pi(\rho-\epsilon)}{(\rho-\epsilon)^{n+1}(\rho-\epsilon-1)} = \frac{M_1}{(\rho-\epsilon)^n(\rho-\epsilon-1)}.$$

Allowing ϵ to approach 0 we have (29.10), and the proof is complete.

The following theorem is an extension of the above result to analytic Jordan curves:

Theorem 29.2. *Let C be an analytic Jordan curve bounding the limited region R. Let $f(z)$ be analytic and uniformly bounded, $|f(z)| \leqq M$, interior to C_ρ, $\rho > 1$. Then we have*

$$\left| f(z) - \sum_0^n a_\nu P_\nu(z) \right| \leqq \frac{AM}{\rho^n}, \qquad z \text{ in } \overline{R}, \tag{29.11}$$

where A is a constant depending only on ρ and C, and $P_\nu(z)$ is the Faber polynomial of degree ν belonging to R.*

Let $z = \psi(w)$ map the exterior of C on the exterior of γ: $|w| = 1$, so that the

* See Faber [1, 2, 3]. It should be pointed out that there is a slight error in the recurrence formula which is given for $P_\nu(z)$ in Faber [3]; formula (15) should contain an additional constant due to the fact that the power series $\mathfrak{P}_n(t)$ in formula (13) may begin with a non-zero term in t to the 0 power.

points at ∞ correspond to each other. Then from the results of Faber [3] we know that $f(z)=\sum_0^\infty a_\nu P_\nu(z)$, z interior to C_ρ, where

$$P_\nu(z) = c^\nu w^\nu(1 + \theta_\nu(z)G), \tag{29.12}$$

G a fixed constant, c the capacity* of R, and $|\theta_\nu(z)| < 1$ for z interior to C_ρ. Also by Faber [3],

$$|a_\nu| = \left| \frac{1}{2\pi i} \int_{|w|=\rho-\epsilon} \frac{f[\psi(\omega)]d\omega}{c^\nu \omega^{\nu+1}} \right| \leqq \frac{1}{2\pi} \frac{M 2\pi(\rho - \epsilon)}{c^\nu(\rho - \epsilon)^{\nu+1}} = \frac{M}{c^\nu(\rho - \epsilon)^\nu},$$

and letting ϵ approach 0 we have

$$|a_\nu| \leqq \frac{M}{c^\nu \rho^\nu}. \tag{29.13}$$

Furthermore from (29.12) we have

$$|P_\nu(z)| \leqq c^\nu[1 + \theta_\nu(z)G], \qquad z \text{ on } C. \tag{29.14}$$

Hence

$$\left| f(z) - \sum_0^n a_\nu P_\nu(z) \right| = \left| \sum_{n+1}^\infty a_\nu P_\nu(z) \right| \leqq \sum_{n+1}^\infty \frac{M[1 + \theta_\nu(z)G]}{\rho^\nu}$$

$$\leqq MM' \sum_{n+1}^\infty \frac{1}{\rho^\nu} \leqq \frac{AM}{\rho^n}, \qquad z \text{ on } C,$$

and the proof is complete.

30. **Lipschitz conditions, generalized derivatives, and degree of approximation.** We start here with the unit circle and prove the following theorem:

THEOREM 30.1. *Let $f(z)$ be analytic in $|z| < \rho > 1$, continuous in $|z| \leqq \rho$, and have a bounded derivative of order $\alpha > 0$ on C_ρ, $|z| = \rho$. Then we have*

$$|f(z) - T_n(z)| \leqq \frac{M \log n}{n^\alpha \rho^n}, \quad |z| \leqq 1, \tag{30.10}$$

where $T_n(z)$ is the sum of the first n terms of the Taylor development of $f(z)$ about $z=0$.

CASE I. $0 < \alpha \leqq 1$. Since $f(z)$ has a bounded derivative of order α it satisfies a Lipschitz condition of order α by Theorem 17.3. Make the transformation $z = \rho y$, then $f(z) \equiv f(\rho y) \equiv \Phi(y)$, and $\Phi(y)$ satisfies a Lipschitz condition of order α on $|y| = 1$. We need the following result due to the author [1]:

* See §7. Faber maps the exterior of C on the interior of the circle whose radius is the reciprocal of c. The relation (7.11) enters in this connection.

LEMMA 30.2. *If $f(z)$ is analytic in $|z|<1$, continuous in $|z|\leqq 1$, and if $f^{(p-1)}(z)$ satisfies a Lipschitz condition of order $\alpha-p+1$, $0<\alpha-p+1\leqq 1$, on $|z|=1$, then we have*

$$|f(z) - T_n(z)| \leqq \frac{M_1 \log n}{n^\alpha}, \quad |z| \leqq 1, \tag{30.11}$$

where $T_n(z)$ is the sum of the first n terms of the Taylor development of $f(z)$ about $z=0$.

By the lemma

$$|\Phi(y) - t_n(y)| \leqq \frac{M_1 \log n}{n^\alpha}, \quad |y| \leqq 1, \tag{30.12}$$

where $t_n(y)$ is the sum of the first n terms of the Taylor development of $\Phi(y)$ about $y=0$. Since $\Phi(y)-t_n(y)$ has a zero of multiplicity at least as great as $n+1$ at $y=0$, we have by Schwarz' lemma*

$$|\Phi(y) - t_n(y)| \leqq \frac{M_1 \log n}{n^\alpha}|y|^{n+1}, \quad |y| \leqq 1, \tag{30.13}$$

and transforming back to the z-plane we have

$$|f(z) - T_n(z)| \leqq \frac{M_1 \log n}{n^\alpha \rho^{n+1}} = \frac{M \log n}{n^\alpha \rho^n}, \quad |z| \leqq 1. \tag{30.14}$$

CASE II. $0\leqq p-1\leqq \alpha<p$. If $f(z)$ has a bounded derivative of order α on C_ρ, then by Theorem 17.3 the function $f^{(p-1)}(z)$ satisfies a Lipschitz condition of order $\alpha-p+1$ on C_ρ. Thus the same method as above yields

$$|f(z) - T_n(z)| \leqq \frac{M \log n}{n^{\alpha-p+1+p-1}\rho^n} = \frac{M \log n}{n^\alpha \rho^n}, \quad |z| \leqq 1, \tag{30.15}$$

and the proof is complete.

As an immediate consequence of a theorem of the author [1] and Theorem 17.3 we have

THEOREM 30.3. *Let C be an analytic Jordan curve and let $f(z)$ be analytic interior to C and continuous in the corresponding closed region bounded by C, and have a bounded derivative of order $\alpha>0$ on C. Then we have*

$$\left| f(z) - \sum_0^n a_\nu P_\nu(z) \right| \leqq \frac{M \log n}{n^\alpha}, \quad z \text{ on } C, \tag{30.16}$$

where $P_\nu(z)$ is the Faber polynomial of degree ν belonging to C.

* I am indebted to the referee for this suggestion, which shortens my original proof considerably.

An application of this theorem leads to

THEOREM 30.4. *Let C be an analytic Jordan curve and let $f(z)$ be analytic interior to C_ρ, $\rho>1$, continuous in the corresponding closed region, and have a bounded derivative of order $\alpha>0$ on C_ρ. Then we have*

$$\left| f(z) - \sum_0^n a_\nu P_\nu(z) \right| \leqq \frac{M \log n}{n^\alpha \rho^n}, \qquad z \text{ on } C, \tag{30.17}$$

where $P_\nu(z)$ is the Faber polynomial of degree ν belonging to C.

Since the Faber development is unique the Faber development of the function

$$\Phi(z) = f(z) - \sum_0^n a_\nu P_\nu(z)$$

is $\sum_{n+1}^\infty a_\nu P_\nu(z)$. Applying Theorem 29.2 to the function $\Phi(z)$ which on C_ρ is in modulus less than $M \log n/n^\alpha$ (by Theorem 30.2), we immediately obtain the desired result.*

31. **The converse problem.** In all of the theorems of §30 we assume conditions on the function $f(z)$ on C_ρ and find the degree of approximation on C. In the following theorem we consider the converse problem:

THEOREM 31.1. *Let E, with boundary C, be a closed limited point set whose complement is simply connected. Let $f(z)$ be analytic on E. If there exists a polynomial $P_n(z)$ of degree n, $n=1, 2, \cdots$, such that*

$$| f(z) - P_n(z) | \leqq \frac{M}{n^{\alpha+1}\rho^n}, \quad \alpha > 0, \quad \rho > 1, \quad z \text{ on } E, \tag{31.10}$$

where M is a constant independent of n and z, then $f(z)$ is analytic interior to C_ρ, continuous in the closed limited region bounded by C_ρ, and has a bounded derivative of every order $\alpha'<\alpha$ on C_ρ.

The analyticity of $f(z)$ interior to C_ρ and its continuity in $\overline{C}_\rho$ are immediate consequences of (31.10). For the bounded derivative we know by hypothesis that

$$| f(z) - P_n(z) | \leqq \frac{M}{n^{\alpha+1}\rho^n}, \qquad z \text{ on } E,$$

$$| f(z) - P_{n+1}(z) | \leqq \frac{M}{(n+1)^{\alpha+1}\rho^{n+1}}, \qquad z \text{ on } E;$$

---

* This proof of the theorem was suggested by the referee.

hence

$$| P_{n+1}(z) - P_n(z) | \leqq \frac{2M}{n^{\alpha+1}\rho^n}, \qquad z \text{ on } E,$$

and consequently by Lemma 19.1 we have

$$| P_{n+1}(z) - P_n(z) | \leqq \frac{2M\rho^{n+1}}{n^{\alpha+1}\rho^n}, \qquad z \text{ on } C_\rho.$$

Since C_ρ is an analytic Jordan curve, we can apply Theorem 22.1 to obtain the inequality

$$\begin{aligned} | P_{n+1}^{\alpha'}(z) - P_n^{\alpha'}(z) | &\leqq \frac{2MK(\alpha', C_\rho)\rho(n+1)^{\alpha'}}{n^{\alpha+1}} \\ &= \frac{M_1}{n^{\alpha-\alpha'+1}}. \end{aligned}$$

By hypothesis $\alpha-\alpha'>0$ and we have uniform convergence, and an application of Theorem 16.1 completes the proof.

In this theorem we assume in the denominator an n to the power $\alpha+1$, whereas in Theorem 30.4 we obtain in the denominator only an n to the power α. This seems to indicate that the hypothesis in Theorem 31.1 is stronger than necessary, but the following example, suggested by Walsh, shows that the exponent $\alpha+1$ cannot be replaced by a smaller exponent:

Let E be the set $|z| \leqq 1$, and C_ρ the circle $|z| = \rho > 1$, and consider the function $f(z) = 1/(\rho - z)$. We know that

$$(31.11) \qquad f(z) - p_n(z) = \frac{1}{\rho}\left[\frac{z^{n+1}}{\rho^{n+1}} + \frac{z^{n+2}}{\rho^{n+2}} + \cdots\right], \qquad z \text{ on } E,$$

where $p_n(z)$ is a polynomial of degree n in z. Integrating both sides of (31.11) from 0 to z, we have

$$(31.12) \qquad F(z) - P_n(z) = \frac{1}{n\rho^n}\left[\frac{nz^{n+2}}{(n+2)\rho} + \frac{nz^{n+3}}{(n+3)\rho^2} + \cdots\right], \qquad z \text{ on } E.$$

The series in brackets converges and consequently

$$| F(z) - P_n(z) | \leqq \frac{M}{n\rho^n}, \qquad z \text{ on } E,$$

and yet the function $F(z) = \log \rho - \log(\rho - z)$ has a logarithmic singularity on C_ρ and its derivative has a pole of the first order on C_ρ. Integrating again we obtain

$$|F^*(z) - P_n^*(z)| \leqq \frac{M}{n^2\rho^n}, \qquad z \text{ on } E.$$

The first derivative of $F^*(z)$ has a logarithmic singularity at $z=\rho$ and hence the function does not have a bounded derivative of order 1 on C_ρ; also we can show that $F(z)$ does not have a bounded derivative of order α, $0<\alpha<1$, on C_ρ, and that $F^*(z)$ does not have a bounded derivative of any order $\alpha \geqq 1$ on C_ρ. On the other hand a direct evaluation shows that $F^*(z)$ has a bounded derivative of any order less than 1 on C_ρ.

This example also shows that the existence of polynomials converging to a function on C with *an error less than* $M/(n\rho^n)$ *does not imply even the continuity of the function on* C_ρ; here $F(z)$ has a logarithmic singularity on C_ρ. We see further that we have an error less than M/ρ^n on C for a function with a single pole of the first order on C_ρ; in this connection the following result on best approximation is of interest:

Let C be an analytic Jordan curve and let $f(z)$ be analytic on and interior to C_ρ except for a finite number of poles on C_ρ as follows: $z_1, z_2, \cdots, z_s$, of orders $h_1, h_2, \cdots, h_s$, respectively, where $h_1 \geqq h_2 \geqq \cdots \geqq h_s > 0$. Let $K_n(z)$ be that polynomial of degree at most n for which

$$\epsilon_n = \max_{z \text{ on } C} |f(z) - K_n(z)|$$

is as small as possible. Then

$$\frac{n^{h_1-1}B_1}{\rho^n} \leqq \epsilon_n \leqq \frac{n^{h_1-1}B_2}{\rho^n},$$

where B_1 and B_2 are constants independent of n and z, so long as z is on C. Faber [3, pp. 105–106] establishes this result in the case of a single pole of the first order on C_ρ and since the method is entirely applicable to the general case the details will not be included here.†

BIBLIOGRAPHY

AHLFORS, L. V.

1. *Untersuchungen zur Theorie der konformen Abbildung und der ganzen Funktionen*, Acta Societatis Scientiarum Fennicae, Nova Series A, vol. 1, (1930), pp. 5–40.

BERNSTEIN, S.

1. *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*, Mémoire de l'Académie Royale de Belgique, (2), vol. 4, pp. 1–104.
2. *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réele*, Paris, 1926.

CARATHÉODORY, C.

1. *Conformal Representation*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 28, 1932.

† For E a segment of the axis of reals see Bernstein [2], Chap. 3.

CURTISS, JOHN.
1. *A note on the degree of polynomial approximation*, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 29–30, Abstract No. 29.
DOETSCH, G.
1. *Probleme aus der Theorie der Wärmteilung, V Mitteilung, Explizite Lösung des Bolometerproblems*, Mathematische Zeitschrift, vol. 28 (1928), pp. 567–578.
FABER, G.
1. *Über polynomische Entwickelungen*, I, Mathematische Annalen, vol. 57 (1903), pp. 389–408.
2. *Über polynomische Entwickelungen*, II, ibid., vol. 64 (1907), pp. 118–135.
3. *Über Tchebyscheffsche Polynome*, Journal für die reine und angewandte Mathematik (Crelle), vol. 150 (1920), pp. 79–106.
4. *Über nach Polynomen fortschreitenden Reihen*, Sitzungsberichte der mathematisch-physikalischen Klasse der Bayerischen Akademie der Wissenschaften zu München, 1922, pp. 157–178.
HARDY, G. H., AND LITTLEWOOD, J. E.
1. *Some properties of fractional integrals*, I, Mathematische Zeitschrift, vol. 27 (1928), pp. 565–606.
2. *Some properties of fractional integrals*, II, ibid., vol. 34 (1932), pp. 403–439.
JACKSON, DUNHAM.
1. *The Theory of Approximation*, American Mathematical Society Colloquium Publications, vol. 11, 1930.
2. *On certain problems of approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 36 (1930), pp. 851–857.
3. *On the application of Markoff's theorem to problems of approximation in the complex domain*, ibid., vol. 37 (1931), pp. 883–890.
KELLOGG, O. D.
1. *Harmonic functions and Green's integral*, these Transactions, vol. 13 (1912), pp. 109–132.
LEVY, P.
1. *Sur la dérivation et l'intégration généralisée*, Bulletin des Sciences Mathématiques de France, vol. 58 (1923), pp. 307–320.
LINDELÖF, E.
1. *Sur un principe général de l'analyse*, Acta Societatis Scientiarum Fennicae, vol. 46, no. 4 (1915), pp. 1–35.
LIOUVILLE, J.
1. *Sur quelques questions de géométrie et de mécanique, et sur un nouveau genre de calcul pour résoudre ces questions*, Journal de l'École Polytechnique, sér. 1, vol. 13, no. 21 (1832), pp. 1–69.
MARKOFF, A.
1. *Sur une question posée par Mendelieff*, Bulletin of the Academy of Sciences of Saint Petersburg, vol. 62 (1889), pp. 1–24.
MONTEL, P.
1. *Sur les polynomes d'approximation*, Bulletin de la Société Mathématique de France, vol. 46 (1919), pp. 151–196.
OSGOOD, W. F., AND TAYLOR, E. H.
1. *Conformal transformations on the boundaries of their regions of definition*, these Transactions, vol. 14 (1913), pp. 277–298.
RIEMANN, B.
1. *Gesammelte Mathematische Werke und Wissenschaftlicher Nachlass*, Leipzig, 1892, pp. 331–344.
RIESZ, M.
1. *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutscher Mathematiker Vereinigung, vol. 23 (1914), pp. 354–368.
2. *Über einen Satz des Herrn Serge Bernstein*, Acta Mathematica, vol. 40 (1916), pp. 337–347.

Seidel, W.

1. *Über die Ränderzuordnung bei konformen Abbildungen*, Mathematische Annalen, vol. 104 (1931), pp. 182–243.

Sewell, W. E.

1. *Degree of approximation by polynomials to continuous functions*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 111–117.
2. *Generalized derivatives and approximation*, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 255–258.
3. *On the modulus of the derivative of a polynomial*, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 699–701.

Szegö, G.

1. *Über trigonometrische und harmonische Polynome*, Mathematische Annalen, vol. 79 (1918), pp. 323–339.
2. *Über einen Satz von A. Markoff*, Mathematische Zeitschrift, vol. 23 (1925), pp. 45–61.
3. *Zur Theorie der schlichten Abbildungen*, Mathematische Annalen, vol. 100 (1928), pp. 188–211.

Tamarkin, J. D.

1. *On integrable solutions of Abel's integral equation*, Annals of Mathematics, (2), vol. 31 (1930), pp. 219–229.

Visser, C.

1. *Über beschränkte analytische Funktionen und die Randverhältnisse bei konformen Abbildungen*, Mathematische Annalen, vol. 105 (1932), pp. 28–39.

Walsh, J. L.

1. *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 20, 1935.

Walsh, J. L., and Sewell, W. E.

1. *Sufficient conditions for various degrees of convergence by polynomials*, Bulletin of the American Mathematical Society, vol. 42 (1936), p. 342, Abstract No. 236.

Warschawski, S.

1. *Über einen Satz von O. D. Kellogg*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1932, pp. 73–86.
2. *Über die Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung*, Mathematische Zeitschrift, vol. 35 (1932), pp. 321–456.

Weyl, H.

1. *Bemerkungen zum Begriff des Differentialquotienten gebrochener Ordnung*, Vierteljahrschrift der Naturforschenden Gesellschaft in Zürich, vol. 62 (1917), pp. 296–302.

Harvard University,
Cambridge, Mass.

STEREOGRAPHIC PARAMETERS AND PSEUDO-MINIMAL HYPERSURFACES, II*

BY

OTTO LAPORTE

Introduction

In a recent paper† G. Y. Rainich and the writer showed that the Weierstrass formulas for minimal surfaces in terms of stereographic parameters can be generalized naturally to hypersurfaces the sum of whose radii of curvature vanishes. The name pseudo-minimal hypersurfaces was proposed for these, since the name minimal should be reserved for the hypersurfaces, the sum of the curvatures of which is zero everywhere.

As was shown in I, a general hypersurface can be represented in terms of stereographic parameters $x_1, \cdots, x_n$ as follows:

$$X_0 = \frac{1}{\lambda}\left(\phi - x_\alpha \frac{\partial \phi}{\partial x_\alpha}\right) \tag{1}$$

$$X_i = \frac{\partial \phi}{\partial x_i} + \frac{x_i}{\lambda}\left(\phi - x_\alpha \frac{\partial \phi}{\partial x_\alpha}\right) \tag{2}$$

with the abbreviation

$$2\lambda = 1 + x_\alpha x_\alpha. \tag{3}$$

Here $X_0, X_1, \cdots, X_n$ are the coordinates of the embedding E_{n+1}, and ϕ is an arbitrary function of $x_1, \cdots, x_n$ whose geometrical meaning is: distance of the tangent plane from the origin multiplied by λ. If now the hypersurface is to be pseudo-minimal, then, as was shown in I, §7, ϕ has to satisfy the following linear partial differential equation:

$$\lambda \frac{\partial^2 \phi}{\partial x_\alpha \partial x_\alpha} - n x_\alpha \frac{\partial \phi}{\partial x_\alpha} + n\phi = 0. \tag{4}$$

In I an infinity of particular solutions of (4) were given. We shall sometimes refer to ϕ as a pseudo-minimal potential.

The present paper contains extensions of the theory in two respects.

1. The Weierstrass formulas for two-dimensional minimal surfaces established a connection between minimal surfaces and analytic functions. Cor-

* Presented to the Society, September 1, 1936; received by the editors April 4, 1936.

† These Transactions, vol. 39 (1936), pp. 154–182. This paper will be referred to as I.

responding to each potential function $\phi(x_1, \cdots, x_n)$ satisfying (4), there belongs a pseudo-minimal hypersurface; but since the latter is a generalization of a minimal surface, we have to regard the functions ϕ satisfying (4) as a generalization of analytic functions. In §4 of I it was shown that the possibility of rotating the $(n+1)$-dimensional sphere upon whose hypersurface our pseudo-minimal hypersurface is mapped results in the invariance of our formulas under the substitutions of a group Ω, whose infinitesimal operators were obtained. This group is the analogue to the displacement group, and it permits therefore to obtain solutions referred to points other than the origin $x_i=0$. Such solutions will be constructed by applying the group transformation especially to the central symmetric solutions of (4). The new solutions will be found to be symmetric in the coordinates of the singularity and of the point under consideration. Upon development of these solutions into series proceeding according to the particular solutions found in I, §§7 and 8, remarkably symmetrical formulas are obtained; they are, for the type of generalized potential function considered here, the analogue of the geometric series and form therefore the prototype for all other expansions.

2. Two-dimensional one-sided minimal surfaces, i.e., surfaces which like a Möbius strip permit a continuous transition from one side to the other, have been the subject of numerous investigations, especially by Lie, Henneberg, and Schilling* and recently by Douglas† from a quite different point of view. The discussion of the possibility of one-sided pseudo-minimal hypersurfaces will be seen to rest upon the behavior of the potential ϕ under a (negative) transformation of reciprocal radii $x_i' = -x_i/r^2$, which transformation is already contained in our group Ω. The result is obtained that in order for a general hypersurface to be one-sided, the potential must satisfy the functional equation

$$\phi(x_1, \cdots, x_n) = -r^2\phi\left(-\frac{x_1}{r^2}, \cdots, -\frac{x_n}{r^2}\right). \tag{5}$$

An infinity of functions satisfying simultaneously (4) and (5) will be constructed, and, thus, an infinity of one-sided pseudo-minimal hypersurfaces will be obtained in spaces of any number of dimensions.

In order not to interrupt the developments of 1 and 2 with derivations of auxiliary formulas pertaining to hypergeometric functions, a third section containing all such formulas has been added.

The author gladly acknowledges his indebtedness to Professor G. Y. Rainich for frequent advice and numerous suggestions.

* See Darboux, *Théorie Générale des Surfaces*, Part 1, Book III, Chap. VI.

† J. Douglas, these Transactions, vol. 34 (1932), p. 731.

1. **Displaced solutions and their expansions.** The particular solutions of the differential equation (4) which we have derived in §§7 and 8 of I were all of the form

$$f_l^{(1)}(r)H_l(x_1, \cdots, x_n) \quad \text{and} \quad f_l^{(2)}(r)H_l(x_1, \cdots, x_n), \tag{1.1}$$

where H_l is a homogeneous polynomial solution of the n-dimensional Laplace equation of degree $l>0$, and where, if F means the Gaussian series,

$$f_l^{(1)}(r) = F\left(l-1, -\frac{n}{2}, l+\frac{n}{2}; -r^2\right), \tag{1.2}$$

$$f_l^{(2)}(r) = r^{2-2l-n}F\left(1-l-n, -\frac{n}{2}, 2-l-\frac{n}{2}; -r^2\right). \tag{1.3}$$

The formulas (3.1), (3.2), and (3.8) of §3 define these functions for any value of the argument. For the cases where these series degenerate into Jacobi polynomials see I, §8.

Like their potential-theoretical analogue: H_l and $r^{-2l-n+2}H_l$, the solutions of the set (1.1) have singularities at $r=0$ and $r=\infty$. It is, however, desirable to know "displaced" solutions which have singularities at an arbitrarily given point. This displacing of the singularity is carried out by an application of the group of the differential equation. Comparing again with the potential case we see that since, for it, the group in question is the group of rigid motions, any harmonic function having a singularity at the origin can immediately be "displaced" by replacing the argument x_i by $x_i-x_i^0$. In our case the situation is essentially more complicated because the group Ω of the differential equation (4) as defined by the substitution

$$x_j' = \frac{1}{D}\left\{s_{j\rho}x_\rho + s_{j0}(1-\lambda)\right\} \tag{1.4}$$

with

$$D = \lambda + s_{00}(1-\lambda) + s_{0\rho}x_\rho \tag{1.5}$$

differs from the displacement group. It expresses the way in which the stereographic parameters change when the surface as a whole is rotated or, which is the same, when the n-dimensional sphere which defines the parameters is rotated. In order to shift a singularity from the points 0 or ∞ to any others, we must therefore carry out a rotation which affects the component ξ_0 of the unit normal vector. Other rotations, which leave ξ_0 unaffected, will result only in rotations in the $x_1, \cdots, x_n$ hyperplane around the origin. We therefore choose the matrix s_{ab} according to a rotation in the X_0, X_1-plane by an

angle η which may be taken as canonical parameter of the group, and put

$$s_{11} = s_{00} = \cos\eta; \qquad s_{10} = -s_{01} = \sin\eta. \tag{1.6}$$

All other $s_{ik}=\delta_{ik}$. Then the transformed solution is according to I, (4.6):

$$\begin{aligned} T\phi = {} & [\lambda + (1-\lambda)\cos\eta - x_1 \sin\eta] \\ & \cdot \phi\left(\frac{(1-\lambda)\sin\eta + x_1\cos\eta}{\lambda + (1-\lambda)\cos\eta - x_1\sin\eta}, \frac{x_2}{\lambda + (1-\lambda)\cos\eta - x_1\sin\eta}, \cdots\right). \end{aligned} \tag{1.7}$$

Since the operator of the infinitesimal transformation

$$Q_1 = (1-\lambda)\frac{\partial}{\partial x_1} + x_1 x_\rho \frac{\partial}{\partial x_\rho} - x_1 \tag{1.8}$$

appears as first coefficient in an expansion of $T\phi$ in powers of η:

$$T\phi = \phi + \eta Q_1\phi + \cdots, \tag{1.9}$$

we can write (1.7) also

$$T\phi = e^{\eta Q_1}\phi. \tag{1.10}$$

Instead of η one may introduce as parameter the radius vector r_I of that point which before the rotation by η was the south pole, which will therefore now contain the singularity. This is done by the substitution

$$\operatorname{tg}\frac{\eta}{2} = r_I, \quad \text{or} \quad e^{i\eta} = \frac{1 + ir_I}{1 - ir_I}. \tag{1.11}$$

It is sufficient to apply the transformation (1.7) to the case of the central symmetric solution; we will get in this way the analogue of the potential solution $|\vec{r}-\vec{r}_I|^{2-n}$. If we write

$$x_1 = r\cos\vartheta, \tag{1.12}$$

where ϑ is the angle between the rays to the two points under consideration, and apply (1.7) to a solution ϕ_0 solely depending upon λ, we obtain

$$\begin{aligned} T\phi_0 = {} & \frac{1}{\lambda_I}\{\lambda_I\lambda + (1-\lambda_I)(1-\lambda) - r_I r\cos\vartheta\} \\ & \cdot \phi_0\left(\frac{\lambda_I\lambda}{\lambda_I\lambda + (1-\lambda_I)(1-\lambda) - r_I r\cos\vartheta}\right). \end{aligned} \tag{1.13}$$

Here in analogy with (3) the abbreviation

$$2\lambda_I = 1 + r_I^2$$

was used. We see thus that except for a constant factor λ_I^{-1} (which is caused

by our using ϕ instead of the generalized Painvin-function; see I, equation (1.5)) $T\phi_0$ is symmetrical in the coordinates of the singular point and of the point under consideration.

It was pointed out in I, §10, that a general solution of the differential equation (4) may be expanded into a series proceeding according to the particular solutions (1.1). We will apply this here to (1.13) and expand $T\phi_0$ which is the simplest solution with a singularity off the origin, into such a series of solutions possessing singularities at zero or infinity. Using the symmetry of the expression for $\lambda_I T\phi_0$ as a function of r_I and of r we may write

$$\lambda_I T\phi_0 = \sum_{l=0}^{\infty} c_l r_I{}^l f_l(r_I) r^l f_l(r) P_l(\cos\vartheta), \tag{1.14}$$

where the coefficient c_l can only be a function of the summation index l. $f_l(r)$ without upper index can be understood to mean some linear combination of $f_l^{(1)}(r)$ and $f_l^{(2)}(r)$; $P_l(\cos\vartheta)$ means that n-dimensional hyperspherical harmonic, which corresponds to axial symmetry, also called the ultraspherical polynomial. In order to show by an example the determination of c_l and of the special form of the radial function $f_l(r)$ we shall from now on specialize to the case of $n=3$. After having familiarized himself with this case, the reader will have no difficulty in deriving corresponding formulas for hypersurfaces of a higher number of dimensions.

The three-dimensional central symmetric solution was derived in §9 of I and found to be

$$\phi_0 = \frac{1}{r} - 6r + r^3. \tag{1.15}$$

Applying (1.13) to this (considered as a function of λ), the following transformed solution is obtained:

$$T\phi_0 = \frac{1}{2\lambda_I} U, \tag{1.16}$$

$$U = \frac{(1-6r_I^2+r_I^4)(1-6r^2+r^4) - 16r_I r(1-r_I^2)(1-r^2)\cos\vartheta - 32r_I^2r^2\sin^2\vartheta}{(1+r_I^2r^2-2r_Ir\cos\vartheta)^{1/2}(r_I^2+r^2+2r_Ir\cos\vartheta)^{1/2}}.$$

We note here, because of the use we shall make of them later on, the following two relations:

$$U(r_I, r, \vartheta) = U(r, r_I, \vartheta), \tag{1.17}$$

$$U\left(r_I, \frac{1}{r}, \vartheta\right) = \frac{1}{r^2} U(r_I, r, \pi - \vartheta). \tag{1.18}$$

As the denominator of (1.16) indicates, the new solution has two singularities, namely, at

$$(1.19) \qquad r = r_I^{-1}, \qquad \vartheta = 0;$$

and

$$(1.20) \qquad r = r_I, \qquad \vartheta = \pi.$$

As could be anticipated from what was said above, these points are the stereographic projections of the former north pole and its antipodal point, the former south pole.

Geometrically the appearance of a singular point P in the potential ϕ of a hypersurface means that for points on the hypersurface infinitely far from the origin $X_\alpha=0$ of the embedding E_{n+1} the projecting vector passes through P. Now in §9 of I several examples of the surfaces belonging to central symmetric potentials were computed, and it was seen that at points infinitely far from the origin the hypersurfaces possess horizontal tangent planes with normal vectors pointing both vertically up and down. (By vertical and horizontal we mean, of course, directions parallel and perpendicular to the X_0-axis.) The stereographic projections of these normal vectors will be the points 0 and ∞ of the $(x_1, \cdots, x_n)$-hyperplane. If the polar axis is inclined by an angle η, these two points will, of course, become the points (1.19) and (1.20).

The appearance of two singularities complicates the determination of the coefficients and proper linear combinations in (1.14), for three regions of convergence will have to be distinguished:

Region A: $r <$ the smaller of r_I and r_I^{-1}.
Region B: r lies between r_I and r_I^{-1}.
Region C: $r >$ the larger of r_I and r_I^{-1}.

We note here that when interchanging r_I and r the inequality for A becomes that for B, and when replacing r by r^{-1} the inequality for A becomes that for C, while that for B is reproduced.

In region A we now write (1.14) with indetermined coefficients

$$(1.21)\ U_A(r_I, r, \vartheta) = \sum_{l=0}^{\infty} c_l r^l [f_l^{(1)}(r) + a_l f_l^{(2)}(r)] r_I^l [f_l^{(2)}(r_I) + b_l f_l^{(1)}(r_I)] P_l(\cos\vartheta);$$

$P_l(\cos\vartheta)$ is now the Legendre polynomial.

$U_B(r_I, r, \vartheta)$ is obtained herefrom by exchanging on the right-hand side r and r_I. We shall see how one can determine the coefficients c_l, a_l, and b_l in an elementary way. Firstly we see that the requirement that U_A remain finite for vanishing r causes us to put

$$a_l = 0. \tag{1.22}$$

Secondly we prove that $b_l=0$ in the following way: In the series for U_B which on account of (1.22) is now

$$U_B(r_I, r, \vartheta) = \sum_{l=0}^{\infty} c_l r_I^l f_l^{(1)}(r_I) r^l [f_l^{(2)}(r) + b_l f_l^{(1)}(r)] P_l(\cos \vartheta) \tag{1.23}$$

we replace r by r^{-1}. The left side changes according to (1.18), while the right-hand series, due to (3.4) and (3.5) becomes

$$\frac{1}{r^2} \sum_{l=0}^{\infty} c_l r_I^l f_l^{(1)}(r_I) r^l \{(-1)^l f_l^{(2)}(r) + b_l[(-1)^{l+1} f_l^{(1)}(r) + \alpha_l f_l^{(2)}(r)]\} P_l(\cos \vartheta).$$

Writing for $\pi-\vartheta$ again ϑ we get as a second form of U_B:

$$U_B(r_I, r, \vartheta) = \sum_{l=0}^{\infty} c_l r_I^l f_l^{(1)}(r_I) r^l \{[1 + (-1)^l \alpha_l b_l] f_l^{(2)}(r) - b_l f_l^{(1)}(r)\} P_l(\cos \vartheta).$$

Comparing this with (1.23) we see that

$$b_l = 0, \tag{1.24}$$

so that (1.21) assumes the form:

$$U_A(r_I, r, \vartheta) = \sum_{l=0}^{\infty} c_l r^l f_l^{(1)}(r) r_I^l f_l^{(2)}(r_I) P_l(\cos \vartheta). \tag{1.25}$$

Thirdly we determine c_l by giving special values to r and r_I. Let q be a number <1. Put for the moment $r=q/r_I$, and let r_I converge to infinity. The inequalities defining region A remain nevertheless fulfilled. From (1.16) we find

$$\lim_{r_I \to \infty} r_I^{-3} U\left(r_I, \frac{q}{r_I}, \vartheta\right) = (1 + q^2 - 2q \cos \vartheta)^{-1/2}.$$

Then we go similarly to the limit on the right-hand side of (1.25) using auxiliary formula (3.6). Thus, we obtain

$$(1 + q^2 - 2q \cos \vartheta)^{-1/2} = \sum_{l=0}^{\infty} c_l (-1)^l q^l P_l(\cos \vartheta). \tag{1.26}$$

Using the well known definition of the Legendre polynomials we find

$$a_l = (-1)^l.$$

Therefore

$$U_A(r_I, r, \vartheta) = \sum_{l=0}^{\infty} (-1)^l r^l f_l^{(1)}(r) r_I^l f_l^{(2)}(r_I) P_l(\cos \vartheta). \tag{1.27}$$

Exchanging r_I with r carries us into B, as was said above. Therefore,

$$(1.28)\qquad U_B(r_I, r, \vartheta) = \sum_{l=0}^{\infty} (-1)^l r_I^l f_l^{(1)}(r_I) r^l f_l^{(2)}(r) P_l(\cos\vartheta).$$

These expansions are quite as simple as their potential-theoretical analogue. The expansion in region C is obtained by replacing in (1.18) r by r^{-1} and using (3.4). Using (1.10), (1.15), and (1.28) we may now return to the canonical parameter and write:

$$(1.29)\qquad \begin{aligned} &e^{\eta Q_1}\left(\frac{1}{r} - 6r + r^3\right) \\ &\quad = \cos^2\frac{\eta}{2} \sum_{l=0}^{\infty} (-1)^l \left(\operatorname{tg}\frac{\eta}{2}\right)^l f_l^{(1)}\left(\operatorname{tg}\frac{\eta}{2}\right) r^l f_l^{(2)}(r) P_l(\cos\vartheta), \end{aligned}$$

where Q_1 is the infinitesimal operator (1.8) of Ω. This relation suggests an entirely different approach which we might have used in arriving at particular solutions of the differential equation (4). By expanding (1.29) on both sides into powers of η we obtain the displaced solution as a linear combination of solutions obtained by an iterated application of Q_1 on ϕ. Each member of this new sequence of particular solutions is, of course, a linear combination of a finite number of solutions (1.1), but this linear relation does not seem to be a simpler one than is already exhibited by (1.29).

An interesting special case of (1.27) is obtained by putting in (1.27) $r_I = 1$, or $\eta = \pi/2$. This results in the spreading of region A and region C to the unit circle, into which region B itself degenerates. Both singularities lie now on the unit circle opposite each other. From (1.16) and (1.27) we have, making use of (3.7) and putting $l = 2j$,

$$(1.30)\qquad \begin{aligned} &\frac{(1 + r^2)^2 - 8r^2\cos^2\vartheta}{[(1 + r^2)^2 - 4r^2\cos^2\vartheta]^{1/2}} \\ &\quad = -\frac{1}{4}\sum_{j=0}^{\infty} \binom{2}{j+1}\binom{2}{2j+2}^{-1} r^{2j} f_{2j}^{(1)}(r) P_{2j}(\cos\vartheta). \end{aligned}$$

2. **One-sided pseudo-minimal hypersurfaces.** If a hypersurface is to be one-sided, then we must be able to reach, starting from an initial n-uple of values $x_1, \cdots, x_n$ with normal unit vector $+\xi_a$ by a continuous change in the stereographic parameters a final n-uple of values $\bar{x}_1, \cdots, \bar{x}_n$ with unit normal vector $-\xi_a$, while the values of the coordinates in the embedding space are the same for the two points:*

* From now on we shall write simply $F(x)$ for any function $F(x_1, \cdots, x_n)$ of the n stereographic parameters.

$$X_a(x) = X_a(\bar{x}) \qquad (a = 0, 1, \cdots, n). \tag{2.1}$$

Let us see first, what the relation between $x_1, \cdots, x_n$ and $\bar{x}_1, \cdots, \bar{x}_n$ has to be, so that

$$\xi_a(x) = -\xi_a(\bar{x}).$$

From the zero component $\xi_0 = (1-\lambda)/\lambda$ we conclude

$$\bar{\lambda} = \frac{\lambda}{r^2}, \tag{2.2}$$

and from the other components $\xi_i = x_i/\lambda$ using (2.2)

$$\bar{x}_i = -\frac{x_i}{r^2}. \tag{2.3}$$

This is a transformation of reciprocal radii together with a reversion of sign. A transformation of this sort is contained in the governing group Ω, as we see upon putting the transformation matrix s_{ab} of (1.4) equal to the negative unit matrix. We therefore know that by the transformation (2.3) a pseudo-minimal potential ϕ is transformed into another pseudo-minimal potential ϕ'.

The transformation (2.3), when introduced into equation (2.1) results in a functional equation between $\phi(x)$ and $\phi(\bar{x})$. The formulas giving X_a in terms of ϕ are given in the introduction, equations (1) and (2). We shall first investigate the zero component of (2.1) which is

$$\frac{1}{\bar{\lambda}}\left(\phi(\bar{x}) - \bar{x}_\rho \frac{\partial\phi(\bar{x})}{\partial\bar{x}_\rho}\right) = \frac{1}{\lambda}\left(\phi(x) - x_\rho \frac{\partial\phi(x)}{\partial x_\rho}\right).$$

By means of (2.3) we now replace the barred parameters throughout by the unbarred ones and obtain, using the abbreviation

$$\phi(x) + r^2\phi\left(-\frac{x}{r^2}\right) = g(x), \tag{2.4}$$

the differential equation:

$$g = x_\rho \frac{\partial g}{\partial x_\rho}. \tag{2.5}$$

Now we take the other components of (2.1):

$$\frac{\partial\phi(\bar{x})}{\partial\bar{x}_i} + \frac{\bar{x}_i}{\bar{\lambda}}\left(\phi(\bar{x}) - \bar{x}_\rho \frac{\partial\phi(\bar{x})}{\partial\bar{x}_\rho}\right) = \frac{\partial\phi(x)}{\partial x_i} + \frac{x_i}{\lambda}\left(\phi(x) - x_\rho \frac{\partial\phi(x)}{\partial x_\rho}\right).$$

This can be written using again abbreviation (2.4) after elementary transformations:

$$\frac{\partial g}{\partial x_i} + \frac{x_i}{\lambda}\left(g - x_\rho \frac{\partial g}{\partial x_\rho}\right) = 0$$

or

$$\frac{\partial g}{\partial x_i} = \frac{x_i}{\lambda - 1} g. \tag{2.6}$$

Introducing (2.6) into (2.5) we see that

$$g = 0. \tag{2.7}$$

We see therefore that for a hypersurface to be one-sided the potential ϕ *must fulfill the functional equation*:

$$\phi(x) = - r^2 \phi\left(- \frac{x}{r^2}\right). \tag{2.8}$$

Returning to pseudo-minimal hypersurfaces, we will take members of the sequence (1.1) and see if simple linear combinations of them satisfy the above functional equation.

Calling a linear combination of $f_l^{(1)}(r)$ and $f_l^{(2)}(r)$ simply $f_l(r)$, we see that because

$$H_l\left(- \frac{x}{r^2}\right) = (-1)^l r^{-2l} H_l(x)$$

equation (2.8) can be fulfilled if

$$f_l\left(\frac{1}{r}\right) = (-1)^{l+1} r^{2(l-1)} f_l(r). \tag{2.9}$$

We shall show that this equation can indeed be satisfied by a proper choice of constants in the $f_l(r)$. As always with discussions concerning properties of the radial functions it is convenient to separate the cases of odd and even dimension number n.

(α) n is an odd integer. Put

$$f_l(r) = f_l^{(1)}(r) + a_l f_l^{(2)}(r),$$

where a_l is constant. Then using (3.4) and (3.5)

$$f_l\left(\frac{1}{r}\right) = (-1)^{l+1} r^{2(l-1)} \left\{ f_l^{(1)}(r) + [\alpha_l(-1)^{l+1} - a_l] f_l^{(2)}(r) \right\}$$

(2.9) will be fulfilled if

$$a_l = \tfrac{1}{2}(-1)^{l+1}\alpha_l$$

or if

$$f_l(r) = f_l^{(1)}(r) + \tfrac{1}{2}(-1)^{l+1}\alpha_l f_l^{(2)}(r). \tag{2.10}$$

(β) n is an even integer. Put, this time,

$$f_l(r) = b_l f_l^{(1)}(r) + f_l^{(2)}(r),$$

where b_l is a constant. Then using (3.9)

$$f_l\left(\frac{1}{r}\right) = b_l r^{2(l-1)} \frac{1}{\beta_l} f_l^{(2)}(r) + \beta_l r^{2(l-1)} f_l^{(1)}(r)$$

(2.9) will be fulfilled if

$$b_l = (-1)^{l+1}\beta_l$$

or if

$$f_l(r) = (-1)^{l+1}\beta_l f_l^{(1)}(r) + f_l^{(2)}(r). \tag{2.11}$$

The formulas (2.10) and (2.11) show how from each pair of particular solutions (1.1) a pseudo-minimal potential may be obtained which will satisfy the functional equation (2.8) and therefore furnish a one-sided hypersurface. More general ϕ's satisfying (2.8) may of course be constructed by combining linearly for different values of l the thus obtained special solutions, or even by infinite series proceeding according to products of type (2.10) or (2.11) multiplied with the proper solid harmonic H_l.

As a special example let us compute the case $n=2$, $l=2$. Formula (2.11) gives

$$f_2(r) = f_2^{(2)}(r) - 3f_2^{(1)}(r),$$

or using (3.1) and (3.8)

$$f_2(r) = \frac{1}{r^4} + \frac{3}{r^2} - 3 - r^2.$$

As harmonic we choose $x_1^2 - x_2^2$, so that

$$\phi = \left(\frac{1}{r^4} + \frac{3}{r^2} - 3 - r^2\right)(x_1^2 - x_2^2).$$

The surface itself becomes, using (1) and (2):

$$X_0 = (x_1^2 - x_2^2)\left(1 + \frac{1}{r^4}\right),$$

$$X_1 = \quad x_1\left(1 - \frac{1}{r^2}\right)\left[1 - \frac{1}{3}(x_1^2 - 3x_2^2)\left(1 + \frac{1}{r^2} + \frac{1}{r^4}\right)\right],$$

$$X_2 = -\, x_2\left(1 - \frac{1}{r^2}\right)\left[1 - \frac{1}{3}(x_2^2 - 3x_1^2)\left(1 + \frac{1}{r^2} + \frac{1}{r^4}\right)\right].$$

This is the famous Henneberg* surface, which was recognized by Lie and by Schilling to be the simplest one-sided minimal surface.

3. **Auxiliary formulas.** Due to the constant appearance of integer and half-integer parameter values it is easier to derive the following formulas from the well known integral representation of the hypergeometric function rather than obtain them by specialization from the general theory.—This integral representation of the radial function $f_l^{(1,2)}$ is

$$\text{const.} \int s^{l-2}(1 - s)^{n/2}(1 + r^2 s)^{n/2} ds.$$

The multiplicative constant will be chosen throughout so that the coefficient of the lowest power of r is unity. The limits or the path of integration have to be selected so that the integrand either resumes its initial value, or vanishes at both limits. As in I the discussion of even and odd dimension numbers is carried out separately. Also we assume l to be $\geqq 2$.

(α) n is an odd integer. The solution I, (8.4) which is regular for $r=0$ is

$$\text{(3.1)} \qquad \begin{cases} f_l^{(1)}(r) = C_l^{(1)} \displaystyle\int_0^1 s^{l-2}(1 - s)^{n/2}(1 + r^2 s)^{n/2} ds, \\ C_l^{(1)} = \dfrac{\Gamma(l + n/2)}{\Gamma(l - 1)\Gamma(n/2 + 1)}, \end{cases}$$

and the singular solution is

$$\text{(3.2)} \qquad \begin{cases} f_l^{(2)}(r) = C_l^{(2)} \displaystyle\oint s^{l-2}(1 - s)^{n/2}(1 + r^2 s)^{n/2} ds, \\ C_l^{(2)} = \tfrac{1}{2}(-1)^l \dfrac{\Gamma(l + n)}{\Gamma(l + n/2 - 1)\Gamma(n/2 + 1)}. \end{cases}$$

The path in this integral is a loop surrounding both branch points $s=1$ and $s=-r^{-2}$. Contracting the path into a loop around $s=\infty$ and applying the

* L. Henneberg, *Über solche Minimalflächen, die eine vorgeschriebene ebene Kurve zur geodätischen Linie haben*, Zürich Dissertation, 1875.

Cauchy theorem immediately gives the polynomial solutions (I, (8.9)). We wish to derive a relation between $f_l^{(1)}(1/r)$, $f_l^{(1)}(r)$ and $f_l^{(2)}(r)$. Replacing in (3.1) r by r^{-1} and introducing as variable

$$t = -\frac{s}{r^2} \tag{3.3}$$

we obtain

$$f_l^{(1)}\left(\frac{1}{r}\right) = (-1)^{l-1}r^{2(l-1)}C_l^{(1)}\int_0^{-1/r^2} t^{l-2}(1-t)^{n/2}(1+r^2t)^{n/2}dt.$$

Combining this again with (3.1) we can obtain an integral from 1 to $-r^{-2}$; but since these are the branch-points of the integrand, such an integral can be transformed into one taken over a contour surrounding the entire branch cut, which is just the path of (3.2). The result is

$$\left\{\begin{aligned} &f_l^{(1)}\left(\frac{1}{r}\right) + (-1)^l r^{2(l-1)}f_l^{(1)}(r) = \alpha_l r^{2(l-1)}f_l^{(2)}(r), \\ &\alpha_l = \frac{\Gamma(l+n/2)\Gamma(l+n/2-1)}{\Gamma(l-1)\Gamma(l+n)}. \end{aligned}\right. \tag{3.4}$$

The corresponding formula for $f_l^{(2)}(r)$ is obtained by replacing in (3.2) r by r^{-1} and introducing the same t as integration variable:

$$f_l^{(2)}\left(\frac{1}{r}\right) = (-1)^l r^{2(l-1)}f_l^{(2)}(r). \tag{3.5}$$

Since, according to I, (8.6a) and the normalization adopted in this paper,

$$\lim_{r\to 0} r^{2l+n-2}f_l^{(2)}(r) = 1,$$

we have on the basis of (3.5)

$$\lim_{r\to 0} r^n f_l^{(2)}\left(\frac{1}{r}\right) = (-1)^l. \tag{3.6}$$

We furthermore note

$$f_l^{(2)}(1) = 0$$

and

$$f_l^{(2)}(1) = (-1)^{(l+n-1)/2}\binom{n/2}{(l+n-1)/2}\binom{n/2}{l+n-1}^{-1}. \tag{3.7}$$

(β) If n is an even number $=2m$, the s-plane is no longer branched. Therefore, while we may keep the definition (3.1) for $f_l^{(1)}$, the singular solution is now to be defined as:

$$\text{(3.8)}\qquad \begin{cases} f_l^{(2)}(r) = \overline{C}_l^{(2)} \displaystyle\int_0^{-1/r^2} s^{l-2}(1-s)^m(1+r^2s)^m ds \\[2ex] \overline{C}_l^{(2)} = (-1)^{l+1}\dfrac{(l+2m-1)!}{(l+m-2)!m!}. \end{cases}$$

The multiplicative constant is as always chosen so that the coefficient of the lowest power of r is unity. Introducing into (3.8) the variable t as defined by (3.3), one obtains easily:

$$\text{(3.9)}\qquad \begin{cases} f_l^{(2)}\left(\dfrac{1}{r}\right) = \beta_l r^{2(l-1)} f_l^{(1)}(r), \\[2ex] \beta_l = \dfrac{(l+2m-1)!(l-2)!}{(l+m-2)!(l+m-1)!}. \end{cases}$$

This formula is the counterpart of (3.4) for spaces of even dimension number.

Added in proof, December 14, 1936. It is perhaps not superfluous to note that the partial differential equation (4) is essentially that of a homogeneous harmonic function of degree one in an E_{n+1}. Nevertheless the general theory of hyperspherical harmonics is of little help in obtaining the above formulas, for in it only the $n+1$ homogeneous polynomials which have no singularities on the hypersphere are considered, while for our purpose all solutions are needed.

University of Michigan,
Ann Arbor, Mich.

ON COMPLEMENTARY MANIFOLDS AND PROJECTIONS IN SPACES L_p AND l_p†

BY

F. J. MURRAY

Introduction. If Λ is a Banach space, $\mathfrak{M}$ a closed linear subset of Λ, then a closed linear subset $\mathfrak{N}$ such that every $f\epsilon\Lambda$ is uniquely expressible as $g+h$, $g\epsilon\mathfrak{M}$, $h\epsilon\mathfrak{N}$, is called a complementary manifold to $\mathfrak{M}$.

In his treatise on linear operations, Banach‡ presents the following two problems ((B), pp. 244–245).

(a) *To every closed linear subset $\mathfrak{M}$ in L_p, $1<p\neq 2$, does there exist a complementary manifold?*

(b) *To every closed linear subset $\mathfrak{M}$ in l_p, $1<p\neq 2$, does there exist a complementary manifold?*

We show in this paper that the answer to both questions is "no."

In Chapter 1 of this paper we show that if a certain limit has the value ∞, then the answer is negative.§ In Chapters 2 and 3, it is proved that this is indeed the case. In the concluding section we discuss the relation of various other problems to (a) and (b).

CHAPTER 1. $C(\mathfrak{M})$ AND $\overline{C}(\Lambda)$

1.1. Let Λ denote a separable space with a p-norm, i.e., Λ is either L_p or l_p or the set of ordered n-tuples of real numbers $\{(a_1, \cdots, a_n)\}$, $l_{p,n}$, with the norm $\|(a_1, \cdots, a_n)\| = (\sum_{i=1}^n |a_i|^p)^{1/p}$. We also let $l_{p,\infty} = l_p$. The notation $p' = 1/(p-1)$, $1/p+1/p'=1$ will be used throughout.

Let $\mathfrak{M}$ be a closed linear manifold in Λ. Let R denote the set of real numbers $0\leqq a\leqq\infty$, and let $r(a, b) = a/(1+a) - b/(1+b)$, $(\infty/(1+\infty)\equiv 1)$. It is easy to see that R with the metric $|r(a, b)|$ is a complete metric space and is homeomorphic to the closed interval $(0, 1)$.

If $\mathfrak{M}$ is a closed linear manifold in Λ, a limited transformation E such that $E\Lambda = \mathfrak{M}$, $E^2 = E$, is said to project Λ on $\mathfrak{M}$.

If E is a limited transformation, we denote by $|E|$ the bound of E.

LEMMA 1.1.1. *Let $\mathfrak{M}$ be a closed linear manifold in Λ. The existence of a*

---

† Presented to the Society, September 5, 1936; received by the editors August 27, 1936.

‡ *Théorie des Opérations Linéaires*, Warsaw, 1932. We shall refer to this work as (B). The quantities considered in the sequel are assumed to be real valued unless explicitly stated to the contrary.

§ This result was obtained while the author was a National Research Fellow at Brown University, Providence, R. I.

complementary manifold $\mathfrak{N}$ *to* $\mathfrak{M}$ *is equivalent to the existence of a projection* E *of* Λ *on* $\mathfrak{M}$.

Suppose $\mathfrak{N}$ exists. Let E be the transformation which is such that $f=g+h$, $g\epsilon\mathfrak{M}$, $h\epsilon\mathfrak{N}$, then $Ef=g$. Owing to the properties of $\mathfrak{N}$, E is single-valued, additive, homogeneous and defined everywhere. Now let $\{f_i\}$ be a sequence, which approaches f and such that if $f_i=g_i+h_i$, $g_i\epsilon\mathfrak{M}$, $h_i\epsilon\mathfrak{N}$ the g_i form a convergent sequence with limit g'. Then g' is $\epsilon\mathfrak{M}$ and the sequence $h_i=f_i-g_i$ also converges to a $h'\epsilon\mathfrak{N}$. By continuity we have $f=g'+h'$. The uniqueness of the resolution of f now implies that $Ef=g'$ or that E is closed. Theorem 7 of (B), Chapter III, p. 41, now implies that E is bounded. Since the range of E is included in $\mathfrak{M}$ and for every $f\epsilon\mathfrak{M}$, $Ef=f$, we see that the range of E is $\mathfrak{M}$ and $E^2=E$ or E is a projection of Λ on $\mathfrak{M}$.

Now suppose E exists. Let $\mathfrak{N}$ be the set of g's in Λ for which $Eg=0$. Since E is limited and linear, $\mathfrak{N}$ is a closed linear manifold. If $f\epsilon\Lambda$, $f=Ef+(1-E)f$ where $Ef\epsilon\mathfrak{M}$ and $(1-E)f\epsilon\mathfrak{N}$, since $E(1-E)f=(E-E^2)f=0$. On the other hand if $h\epsilon\mathfrak{M}$, $h=Ef$ for some $f\epsilon\Lambda$, and hence $Eh=E^2f=Ef=h$. Thus if $h\epsilon\mathfrak{N}\cdot\mathfrak{M}$, $0=Eh=h$, or $\mathfrak{M}\cdot\mathfrak{N}=\{0\}$. Now let f again be $\epsilon\Lambda$, $f=g+h=g'+h'$, g, $g'\epsilon\mathfrak{M}$, h, $h'\epsilon\mathfrak{N}$. Then $g-g'=h'-h$, and since $h-h'\epsilon\mathfrak{N}$, $g'-g\epsilon\mathfrak{M}$, and $\mathfrak{M}\cdot\mathfrak{N}=\{0\}$, this implies $g-g'=h'-h=0$. This shows that $f\epsilon\Lambda$ can only be expressed in one way as $h+g$, $h\epsilon\mathfrak{N}$, $g\epsilon\mathfrak{M}$.

We may therefore consider problems (a) and (b) in the following equivalent form.

(A) *To every closed linear manifold* $\mathfrak{M}$ *of* L_p, $1<p\neq 2$, *is there a projection of* L_p *on* $\mathfrak{M}$?

(B) *To every closed linear manifold* $\mathfrak{M}$ *of* l_p, $1<p\neq 2$, *is there a projection of* l_p *on* $\mathfrak{M}$?

Let $\Lambda_1, \cdots, \Lambda_n$, $n=1, 2, \cdots, \infty$, $\Lambda_\infty\equiv\Lambda$ be a set of spaces. Let $\sum_{\alpha=1}^n \oplus\Lambda_\alpha=\Lambda_1\oplus\cdots\oplus\Lambda_n$ denote the space of ordered sets of elements $\{f_1, f_2, \cdots, f_n\}$ $(f_\infty\equiv f)$ $f_\alpha\epsilon\Lambda_\alpha$, such that $\sum_{\alpha=1}^n\|f_\alpha\|^p<\infty$, with a norm defined by the equation

$$\|\{f_1, f_2, \cdots, f_n\}\| = \left(\sum_{\alpha=1}^n \|f_\alpha\|^p\right)^{1/p}.$$

$\Lambda_\alpha\cong\Lambda_\beta$ is to mean that there exists a one-to-one isometric mapping of Λ_α on Λ_β.

LEMMA 1.1.2. (a) $\sum_{\alpha=1}^n\oplus l_{p,n_\alpha}=l_{p,m}$, $n_\alpha=1, 2, \cdots, \infty$, *if* $\sum_{\alpha=1}^n n_\alpha=m$.

(b) $\sum_{\alpha=1}^n\oplus\Lambda_\alpha\cong L_p$ *if* $\Lambda_\alpha=L_p$, *for each* α.

The proof of this lemma may be left to the reader.

1.2. Let $\mathfrak{M}$ be a closed linear manifold in Λ. We define a function $C(\mathfrak{M})$,

which takes on values in R as follows. If there exists no projection of Λ on $\mathfrak{M}$, then $C(\mathfrak{M})=\infty$. Otherwise $C(\mathfrak{M})=$ g.l.b. $(|E|; E\Lambda=\mathfrak{M}, E^2=E)$. Similarly we define the function $\overline{C}(\Lambda)$ as l.u.b. $(C(\mathfrak{M}), \mathfrak{M}\subseteq\Lambda)$.

LEMMA 1.2.1. *Let Λ_1 and Λ_2 be such that Λ_1 is equivalent* [(B), p. 180] *to a closed linear manifold $\mathfrak{M}$ of Λ_2. Let $\mathfrak{M}$ be such that there exists a projection E of Λ_2 on $\mathfrak{M}$, with $|E|=1$, $\mathfrak{N}$ the set of f's $\epsilon\Lambda$, for which $Ef=0$. Let $\mathfrak{P}$ be any closed linear manifold of Λ_2, such that if $f\epsilon\mathfrak{P}$, then $f=g+h$, $g\epsilon\mathfrak{P}\cdot\mathfrak{M}$, $h\epsilon\mathfrak{P}\cdot\mathfrak{N}$. Let $\mathfrak{P}_1$ in Λ_1 be the manifold which corresponds to $\mathfrak{P}\cdot\mathfrak{M}$. Then $C(\mathfrak{P}_1)\leqq C(\mathfrak{P})$.*

If $C(\mathfrak{P})=\infty$, our statement is true. Suppose $C(\mathfrak{P})$ is $<\infty$. Let F be any projection of Λ_2 on $\mathfrak{P}$. Then EF is a projection on $\mathfrak{P}\cdot\mathfrak{M}$. For if $f_1\epsilon\Lambda_2$, $f=Ff_1=g+h$, $g\epsilon\mathfrak{P}\cdot\mathfrak{M}$, $h\epsilon\mathfrak{P}\cdot\mathfrak{N}$, and $EFf=g$ or the range of EF is included in $\mathfrak{P}\cdot\mathfrak{M}$. Also for every $h\epsilon\mathfrak{P}\cdot\mathfrak{M}$, we have $EFh=Eh=h$. This with our previous statement shows that $(EF)^2=EF$ and that the range of EF is exactly $\mathfrak{P}\cdot\mathfrak{M}$.

Let $(EF)'$ be EF considered only on $\mathfrak{M}$. Obviously $(EF)'$ projects $\mathfrak{M}$ on $\mathfrak{P}\cdot\mathfrak{M}$. Let G be the corresponding transformation on Λ_1. Then $C(\mathfrak{P}_1)\leqq|G|=|(EF)'|\leqq|EF|\leqq|E|\cdot|F|=|F|$ or $C(\mathfrak{P}_1)\leqq|F|$. Since F was any projection on $\mathfrak{P}$, $C(\mathfrak{P}_1)\leqq C(\mathfrak{P})$.

LEMMA 1.2.2. *If Λ_1 and Λ_2 are as in Lemma* 1.2.1, $\overline{C}(\Lambda_1)\leqq\overline{C}(\Lambda_2)$. *In particular if $\Lambda_2=\Lambda_0\oplus\Lambda_1$, $\overline{C}(\Lambda_1)\leqq\overline{C}(\Lambda_2)$.*

Let $\mathfrak{P}_1$ be any closed linear manifold of Λ_1, $\mathfrak{P}$ the corresponding set of elements in $\mathfrak{M}$. $\mathfrak{P}$ is a closed linear manifold satisfying the conditions given in Lemma 1.2.1, since $\mathfrak{P}\cdot\mathfrak{M}=\mathfrak{P}$, $\mathfrak{P}\cdot\mathfrak{N}=\{0\}$. Lemma 1.2.1 now implies that $C(\mathfrak{P}_1)\leqq C(\mathfrak{P})\leqq\overline{C}(\Lambda_2)$. But $\mathfrak{P}_1$ was any closed linear manifold in Λ_1, hence $\overline{C}(\Lambda_1)\leqq\overline{C}(\Lambda_2)$.

To show the second statement, we take $\mathfrak{M}\subseteq\Lambda_0\oplus\Lambda_1$ as the set of elements $\{0,f\}$ of $\Lambda_0\oplus\Lambda_1$, E as the transformation of $\Lambda_0\oplus\Lambda_1$, such that $E\{f,g\}=\{0,g\}$. One readily sees that $\mathfrak{M}$ is equivalent to Λ_1 and that E projects $\Lambda_0\oplus\Lambda_1$ on $\mathfrak{M}$ and $|E|=1$. We may now apply the first part of this lemma to obtain the desired result.

LEMMA 1.2.3. *If $\Lambda\cong\sum_{\alpha=1}^{\infty}\oplus\Lambda_\alpha$ and k is* lim $\sup_{\alpha\to\infty}\overline{C}(\Lambda_\alpha)$, *then there exists a manifold $\mathfrak{P}\subseteq\Lambda$, such that $C(\mathfrak{P})\geqq k$.*

It follows from the definition of k, that if ϵ is >0, then there exists an infinite number of the α's for which $r(\overline{C}(\Lambda_\alpha), k)\geqq-\epsilon$. Thus we can find a sequence of integers $\{\alpha_i\}$ such that $\alpha_i<\alpha_{i+1}$, for which $r(\overline{C}(\Lambda_{\alpha_i}), k)\geqq-2^{-i-1}$.

Now since $r(\overline{C}(\Lambda_{\alpha_i}), k)\geqq-2^{-i-1}$, we can find a $\mathfrak{P}_{\alpha i}$, in $\Lambda_{\alpha i}$, such that $r(C(\mathfrak{P}_{\alpha i}), k)>-2^{-i}$. Let $\mathfrak{P}$ be the closed linear manifold consisting of those elements $\{f_1, f_2, f_3, \cdots\}\epsilon\Lambda$, such that $f_\beta=0$ if β is not $\epsilon\{\alpha_i\}$ and $f_{\alpha_i}\epsilon\mathfrak{P}_{\alpha_i}$. As we saw in the proof of Lemma 1 2.2, Λ_{α_i} and Λ are as Λ_1 and Λ_2 in Lemma 1.2.1

and it is easily seen that $\mathfrak{P}$ satisfies the conditions given in Lemma 1.2.1 also. Thus Lemma 1.2.1 now implies that $C(\mathfrak{P})$ is $\geqq C(\mathfrak{P}_{\alpha_i})$. Hence $r(C(\mathfrak{P}), k) \geqq -2^{-i}$ for every i. This implies that $r(C(\mathfrak{P}), k) \geqq 0$, $C(\mathfrak{P}) \geqq k$.

1.3. We now prove the following lemma.

Lemma 1.3.1. $\overline{C}(L_p) \geqq \overline{C}(l_{p,\infty})$.

In (B), Theorem 9, Chapter XII, p. 206, it is shown that the manifold $\mathfrak{M} \subseteq L_p$, determined by the functions y_i is equivalent to l_p when

$$y_i(t) = 2^{i/p} \text{ for } 1/2^i \leqq t \leqq 1/2^{i-1}, \quad y_i(t) = 0, \quad \text{otherwise}.$$

Now for any $z(t) \in L_p$, let

$$E(z(t)) = \sum_{i=1}^{\infty} \int_0^1 z(s) y_i^{p-1}(s) ds \cdot y_i(t).$$

Then by a direct calculation one can verify that $|E| = 1$ and that if $z \in \mathfrak{M}$ (i.e., if $z = \sum_i \alpha_i y_i$, $\sum |\alpha_i|^p < \infty$), then $Ez = z$. Hence E projects L_p on $\mathfrak{M}$ and we may apply Lemma 1.2.2 so that it yields $\overline{C}(L_p) \geqq \overline{C}(l_{p,\infty})$.

Lemma 1.3.2. *There exists a linear manifold $\mathfrak{M} \subseteq L_p$, such that $C(\mathfrak{M}) = \overline{C}(L_p)$.*

This follows from Lemma 1.1.2, (b) (with $n = \infty$) and Lemma 1.2.3 for k is in this case $\overline{C}(L_p)$.

Lemma 1.3.3. *There exists a linear manifold $\mathfrak{M} \subseteq l_{p,\infty}$, such that $C(\mathfrak{M}) = \overline{C}(l_{p,\infty})$.*

In Lemma 1.1.2, (a), let $n_\alpha = \infty$ for every α. Then apply Lemma 1.2.3.

Lemma 1.3.4. $\overline{C}(l_{p,n}) \geqq \overline{C}(l_{p,m})$ *if* $n \geqq m$.

This follows from Lemma 1.1.2, (a) and Lemma 1.2.2.

Theorem I. *$C(\mathfrak{M})$ and $\overline{C}(\Lambda)$ are to be as in §1.2. There exists an $\mathfrak{M}$ in L_p, and an $\mathfrak{N}$ in l_p, such that $C(\mathfrak{M}) = \overline{C}(L_p)$ and $C(\mathfrak{N}) = \overline{C}(l_p)$. Furthermore*

$$1 = \overline{C}(l_{p,1}) \leqq \overline{C}(l_{p,2}) \leqq \cdots \leqq \overline{C}(l_p) \leqq \overline{C}(L_p).$$

The lemmas of this section imply this theorem.

Now if we are able to show that $\lim_{n\to\infty} \overline{C}(l_{p,n}) = \infty$, it follows from this theorem that $\overline{C}(L_p) = \overline{C}(l_p) = \infty$ and then in each of them we have a manifold $\mathfrak{M}$ for which $C(\mathfrak{M}) = \infty$. Hence from the definition of $C(\mathfrak{M})$, we can answer problems (a) and (b) negatively. The next two chapters of this paper contain the proof of the fact that $\lim_{n\to\infty} \overline{C}(l_{p,n}) = \infty$.

Chapter 2. $\mathfrak{M}$ in Situation A

2.1. Let $f = \{a_1, \cdots, a_n\}$ be an n-dimensional vector, which may be regarded as $\in l_{p,n}$. We define for $k > 0$

$$\{f\}^k = \{ |a_1|^k \operatorname{sign} a_1, \cdots, |a_n|^k \operatorname{sign} a_n\}$$

$$[f]^k = \{ |a_1|^k, \cdots, |a_n|^k\},$$

which may be regarded as elements of $l_{p/k,n}$. If $g=\{b_1, \cdots, b_n\}$, we define $(f, g)=\sum_{i=1}^n a_i b_i$. The linearity and homogeneity of this expression will be used without comment.

The following two lemmas can be easily shown.

LEMMA 2.1.1. *If* $p>1$,

$$\frac{d}{dt}\|f+tg\|^p\Big]_{t=0} = p(\{f\}^{p-1}, g).$$

LEMMA 2.1.2. *If* $p>2$,

$$\frac{d^2}{dt^2}\|f+tg\|^p = p(p-1)([f+tg]^{p-2}, [g]^2).$$

We now prove

LEMMA 2.1.3. *If* $p>2$, *and* $(\{f\}^{p-1}, g)\geqq 0$, *then* $\|f+g\|^p \geqq \|f\|^p$.

By Lemma 2.1.2, $H(t)=\|f+tg\|^p$ is convex in t and hence increasing for $t\geqq 0$ since $dH/dt|_{t=0}\geqq 0$ by Lemma 2.1.1.

2.2. If $\mathfrak{M}$ is a linear manifold in $l_{p,n}$, let $\mathfrak{M}^\perp$ consist of those elements $g\epsilon l_{p',n}$, $1/p+1/p'=1$, for which $(f, g)=0$ for all $f\epsilon\mathfrak{M}$. If $\mathfrak{M}$ is k-dimensional, it is well known that $\mathfrak{M}^\perp$ is $(n-k)$-dimensional and also that $(\mathfrak{M}^\perp)^\perp=\mathfrak{M}$. The following lemma is of a standard type in the theory of linear manifolds of a finite number of dimensions and the proof of it may be omitted.

LEMMA 2.2.1. *Let* $\mathfrak{M}$ *be a k-dimensional linear manifold in* $l_{p,n}$. *Let* $\phi_1, \cdots, \phi_k$ *be k linearly independent elements of* $\mathfrak{M}$. *If E is a projection of* $l_{p,n}$ *on* $\mathfrak{M}$, *there exist k elements* $\psi_1, \cdots, \psi_k$, *of* $l_{p',n}$ *such that for every* $f\epsilon l_{p,n}$,

$$(\alpha) \qquad Ef = \sum_{i=1}^{k} (\psi_i, f)\phi_i$$

and $(\psi_i, \phi_j)=\delta_{i,j}$. *If* ψ_i*'s with this last property are given, the E defined by* (α) *is a projection. If E' is any other projection of* $l_{p,n}$ *on* $\mathfrak{M}$, *then* $\psi_i' = \psi_i + g_i$, $i=1, \cdots, k$, *where* g_i *is* $\epsilon\mathfrak{M}^\perp$.

2.3. If E is a linear transformation in $l_{p,n}$, we denote by E^* (the adjoint of E) the transformation in $l_{p',n}$, such that if g and $g^*\epsilon l_{p',n}$, are related so that for every $f\epsilon l_{p,n}$, $(Ef, g)=(f, g^*)$, then $E^*g=g^*$. It is well known that E^* is linear, $|E|=|E^*|$ and $(FE)^*=E^*F^*$.

LEMMA 2.3.1. *If $\mathfrak{M}$ and E are as in Lemma 2.2.1, then*

$$E^*g = \sum_{i=1}^{k} (\phi_i, g)\psi_i$$

for all $g \epsilon l_{p',n}$.

For every $f \epsilon l_{p,n}$, we have

$$(Ef, g) = \left(\sum_{i=1}^{k} (\psi_i, f)\phi_i, g \right) = \sum_{i=1}^{k} (\psi_i, f)(\phi_i, g) = \left(f, \sum_{i=1}^{k} (\phi_i, g)\psi_i \right).$$

LEMMA 2.3.2. *If $\mathfrak{M}$ and E are as in Lemma 2.2.1, then $1-E^*$ is a projection on $\mathfrak{M}^\perp$.*

The range of $1-E^*$ is $\mathfrak{M}^\perp$. For if $f=(1-E^*)g$ and h is $\epsilon\mathfrak{M}$, then $Eh=h$ and

$$(h, f) = (h, (1 - E^*)g) = (h, g) - (h, E^*g) = (Eh, g) - (h, E^*g) = 0.$$

Hence f is $\epsilon\mathfrak{M}^\perp$ or the range of $1-E^*$ is included in $\mathfrak{M}^\perp$. Furthermore by Lemma 2.3.1, if f is $\epsilon\mathfrak{M}^\perp$,

$$E^*f = \sum_{i=1}^{k} (\phi_i, f)\psi_i = 0$$

and $(1-E^*)f=f$. This with our previous result shows that the range $1-E^*$ is exactly $\mathfrak{M}^\perp$ and $(1-E^*)^2=1-E^*$.

LEMMA 2.3.3. $\overline{C}(l_{p',n}) \leqq \overline{C}(l_{p,n})+1$.

Let $\mathfrak{M}'$ be any linear manifold of $l_{p',n}$. Let $\mathfrak{M}=\mathfrak{M}'^\perp$. Then if ϵ is >0, there exists a projection E of $l_{p,n}$ on $\mathfrak{M}$ with $|E| \leqq C(\mathfrak{M})+\epsilon \leqq \overline{C}(l_{p,n})+\epsilon$. By Lemma 2.3.2, $1-E^*$ is a projection on $\mathfrak{M}^\perp=(\mathfrak{M}'^\perp)^\perp=\mathfrak{M}'$. Thus $C(\mathfrak{M}') \leqq |1-E^*| \leqq 1+|E^*| = 1+|E| \leqq 1+\overline{C}(l_{p,n})+\epsilon$, which implies our lemma.

Of course p and p' are interchangeable and so we see that the answer to our question is the same for both p and p'. Thus we may confine ourselves to the case $p<2$. This is not an essential step in our proof but merely a convenient one. We suppose from now on that p is <2.

2.4. We say that Situation A holds in a k-dimensional manifold $\mathfrak{M}$ of $l_{p,n}$, if

(a) we have k linearly independent elements, $\phi_1, \cdots, \phi_k, \epsilon\mathfrak{M}$, and k elements of $l_{p',n}, \psi_1, \cdots, \psi_k$ such that $(\phi_i, \psi_j)=\delta_{i,j}$ (the transformation E given by the equation $Ef=\sum_{i=1}^{k}(\psi_i, f)\phi_i$ is a projection of $l_{p,n}$ on $\mathfrak{M}$);

(b) we have r elements $h_1, \cdots, h_r$ of $\mathfrak{M}$, with $\|h_i\|=1$;

(c) there exists a constant $C>1$, such that $\|E^*\{h_i\}^{p-1}\|=C$ for every i;

(d) there exist r constants $c_1, \cdots, c_r, c_i>0$, such that for every $f\epsilon\mathfrak{M}$ and $g\epsilon\mathfrak{M}^\perp$

$$\sum_{i=1}^{r} c_i(\{h_i\}^{p-1}, f)(\{E^*\{h_i\}^{p-1}\}^{p'-1}, g) = 0.$$

LEMMA 2.4.1. *If $\mathfrak{M}$ is in Situation* A, *then* $C(\mathfrak{M}) \geqq C$ (cf. (c) above).

Since $|E| = |E^*|$, we must show that for every projection E' of $l_{p,n}$ on $\mathfrak{M}$, $|E'^*| \geqq C$. Since $\|h_i\| = 1$ and hence $\|\{h_i\}^{p-1}\| = 1$, it will be sufficient to show that $\|E'^*\{h_i\}^{p-1}\| \geqq C$ for at least one i.

Now

$$E'^*\{h_i\}^{p-1} = \sum_{j=1}^{k} (\{h_i\}^{p-1}, \phi_j)\psi_j',$$

where $\psi_j' = \psi_j + g_j$, where g_j is an element of $\mathfrak{M}^\perp$ (Lemmas 2.2.1 and 2.3.1). Let E_t be the projection of $l_{p,n}$ on $\mathfrak{M}$ given by

$$E_t f = \sum_{i=1}^{k} (\psi_i + tg_i, f)\phi_i.$$

By Lemma 2.3.1,

$$E_t^* g = \sum_{i=1}^{k} (\phi_i, g)(\psi_i + tg_i)$$

and

$$\begin{aligned} E_t^*\{h_i\}^{p-1} &= \sum_{j=1}^{k} (\phi_j, \{h_i\}^{p-1})(\psi_j + tg_j) \\ &= \sum_{j=1}^{k} (\phi_j, \{h_i\}^{p-1})\psi_j + t\sum_{j=1}^{k} (\phi_j, \{h_i\}^{p-1})g_j \\ &= E^*\{h_i\}^{p-1} + t\sum_{j=1}^{k} (\phi_j, \{h_i\}^{p-1})g_j. \end{aligned}$$

Now by Lemma 2.1.1

$$\begin{aligned} \frac{d}{dt}\|E_t^*\{h_i\}^{p-1}\|_{t=0}^{p'} &= \left(\{E^*\{h_i\}^{p-1}\}^{p'-1}, \sum_{j=1}^{k} (\phi_j, \{h_i\}^{p-1})g_j\right) \\ &= \sum_{j=1}^{k} (\{E^*\{h_i\}^{p-1}\}^{p'-1}, g_j)(\phi_j, \{h_i\}^{p-1}). \end{aligned}$$

Since $g_j \epsilon \mathfrak{M}^\perp$, $\phi_j \epsilon \mathfrak{M}$, (d) implies that

$$\sum_{i=1}^{r} c_i \frac{d}{dt}\|E_t^*\{h_i\}^{p-1}\|^{p'}]_{t=0} = \sum_{i=1}^{r} c_i \sum_{j=1}^{k} (\{E^*\{h_i\}^{p-1}\}^{p'-1}, g_j)(\phi_j, \{h_i\}^{p-1}) = 0.$$

Since $c_i>0$, for $i=1, \cdots, r$ this implies that there must be an i' such that $d\|E_i^*\{h_{i'}\}^{p-1}\|^{p'}/dt]_{t=0}$ is $\geqq 0$. Hence by the above $d\|E^*\{h_{i'}\}^{p-1}+tg'_{i'}\|^{p'}/dt]_{t=0}$ is $\geqq 0$, when $g'_{i'}=\sum_{i=1}^{k}(\phi_i, \{h_{i'}\}^{p-1})g_i$. Lemma 2.1.3 now yields

$$\|E^*\{h_{i'}\}^{p-1}+ g'_{i'}\|^{p'} \geqq \|E^*\{h_{i'}\}^{p-1}\|^{p'}$$

since $p' > 2$. But

$$E'^*\{h_{i'}\}^{p-1} = E_1^*\{h_{i'}\}^{p-1}=E^*\{h_{i'}\}^{p-1}+g'_{i'}$$

and $\|E^*\{h_{i'}\}^{p-1}\|^{p'}=C$. Substituting these values on both sides of our inequality, we get $\|E'^*\{h_{i'}\}^{p-1}\|^{p'}\geqq C^{p'}$ or $\|E'^*\{h_{i'}\}^{p-1}\|\geqq C$. As we remarked at the beginning of the proof this is sufficient.

Chapter 3. The product of $l_{p,n}$ and $l_{p,m}$

3.1. We define $l_{p,n}\otimes l_{p,m}$ as $l_{p,nm}$. If $f=\{a_1, \cdots, a_n\}\,\epsilon\, l_{p,n}$, and $g=\{b_1, \cdots, b_m\}\,\epsilon\, l_{p,m}$, we define $f\otimes g$ as $\{a_1b_1, a_1b_2, \cdots, a_1b_m, a_2b_1, a_2b_2, \cdots, a_2b_m, \cdots, a_nb_1, a_nb_2, \cdots, a_nb_m\}$ or if $f\otimes g=\{c_1, \cdots, c_{nm}\}$, then $c_{(s-1)m+t}=a_sb_t$. The proofs of the following Lemmas 3.1.1–3.1.4 do not present any difficulty and may be left to the reader.

Lemma 3.1.1. (i) $\|f\otimes g\|=\|f\|\cdot\|g\|$,
(ii) $(\phi_1\otimes\phi_2, f\otimes g)=(\phi_1, f)(\phi_2, g)$,
(iii) $\{f\otimes g\}^k=\{f\}^k\otimes\{g\}^k$,
(iv) $\alpha(f^{(1)}\otimes g)+\beta(f^{(2)}\otimes g)=(\alpha f^{(1)}+\beta f^{(2)})\otimes g$.

Lemma 3.1.2. *Let $f_r=\{a_{1,r}, \cdots, a_{n,r}\}$, $r=1, \cdots, k$, $k\leqq n$, be k linearly independent elements of $l_{p,n}$, and $g=\{b_{1,r}, \cdots, b_{m,r}\}$ be k elements of $l_{p,m}$, such that $\sum_{r=1}^{k}f_r\otimes g_r=0$. Then $g_r=0$, $r=1, \cdots, k$.*

Lemma 3.1.3. *Let f_r, $r=1, \cdots, k$, $k\leqq n$, be k linearly independent elements of $l_{p,n}$ and for every $r=1, \cdots, k$ let $g_{r,s}$, $s=1, \cdots, k_r$, $k_r\leqq m$, be k_r linearly independent elements of $l_{p,m}$. Then the set of elements $f_r\otimes g_{r,s}$, $r=1, \cdots, k$, $s=1, \cdots, k$ are linearly independent.*

Lemma 3.1.4. *Let $f_1, \cdots, f_n$ be n linearly independent elements of $l_{p,n}$, $g_1, \cdots, g_m$, m linearly independent elements $l_{p,m}$. Then the set of elements $f_i\otimes g_j$, $i=1, \cdots, n$, $j=1, \cdots, m$, determine $l_{p,mn}$.*

3.2. Let $\mathfrak{M}^{(1)}\subseteq l_{p,n}$ and $\mathfrak{M}^{(2)}\subseteq l_{p,m}$ be linear manifolds. We define $\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)}$ as the linear manifold in $l_{p,mn}$ determined by the elements $f\otimes g$, $f\,\epsilon\,\mathfrak{M}^{(1)}$, $g\,\epsilon\,\mathfrak{M}^{(2)}$.

Lemma 3.2.1. *If $\phi_1^{(1)}, \cdots, \phi_{k^{(1)}}^{(1)}$ is a set of linearly independent elements which determine $\mathfrak{M}^{(1)}$ and $\phi_1^{(2)}, \cdots, \phi_{k^{(2)}}^{(2)}$ is a set of linearly independent elements which determine $\mathfrak{M}^{(2)}$, then $\phi_i^{(1)}\otimes\phi_j^{(2)}$, $i=1, \cdots, k^{(1)}$, $j=1, \cdots, k^{(2)}$, determine the manifold $\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)}$ which is $k^{(1)}k^{(2)}$-dimensional.*

The proof is easily derived from Lemma 3.1.3.

LEMMA 3.2.2. *Let $\mathfrak{M}^{(1)}$ and $\mathfrak{M}^{(2)}$ be as above. Then $(\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)})^\perp$ is determined by the elements of the form $f\otimes g$, where either $f\epsilon\mathfrak{M}^{(1)\perp}$ or $g\epsilon\mathfrak{M}^{(2)\perp}$.*

We first show that if $f\otimes g$ is such that either $f\epsilon\mathfrak{M}^{(1)\perp}$ or $g\epsilon\mathfrak{M}^{(2)\perp}$, then $f\otimes g\epsilon(\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)})^\perp$. Indeed by the definition of $\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)}$ and the linearity of the operation $(\quad,\quad)$, if $(\phi^{(1)}\otimes\phi^{(2)}, f\otimes g)=0$ for all $\phi^{(1)}\epsilon\mathfrak{M}^{(2)}$ and $\phi^{(2)}\epsilon\mathfrak{M}^{(2)}$, then $f\otimes g\epsilon(\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)})^\perp$. But Lemma 3.1.1, (ii) implies that $(\phi^{(1)}\otimes\phi^{(2)}, f\otimes g)=(\phi^{(1)}, f)(\phi^{(2)}, g)$. Since either $(\phi^{(1)}, f)=0$ or $(\phi^{(2)}, g)=0$, we obtain that $f\otimes g\epsilon(\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)})^\perp$.

Now let $\tilde{\phi}_1^{(1)}, \cdots, \tilde{\phi}_{n-k^{(1)}}^{(1)}$ and $\tilde{\phi}_1^{(2)}, \cdots, \tilde{\phi}_{n-k^{(2)}}^{(2)}$ be sets of $n-k^{(1)}$ and $m-k^{(2)}$ linearly independent elements of $\mathfrak{M}^{(1)\perp}$ and $\mathfrak{M}^{(2)\perp}$ respectively. Let $\psi_1^{(1)}, \cdots, \psi_{k^{(1)}}^{(1)}$ $(\psi_1^{(2)}, \cdots, \psi_{k^{(2)}}^{(2)})$ be elements of $l_{p',n}$ $(l_{p',m})$ such that $\tilde{\phi}_i^{(1)}$'s and $\psi_j^{(1)}$'s together determine $l_{p',n}$ ($\tilde{\phi}_i^{(2)}$'s and $\psi_j^{(2)}$'s determine $l_{p',m}$). In Lemma 3.1.3, let $f_r=\psi_r^{(1)}$ for $r=1, \cdots, k^{(1)}$, $f_{k^{(1)}+t}=\tilde{\phi}_t^{(1)}$ for $t=1, \cdots, n-k^{(1)}$; $k=n$. For $r=1, \cdots, k^{(1)}$, let $k_r=m-k^{(2)}$, $g_{r,s}=\tilde{\phi}_s^{(2)}$, $s=1, \cdots, m-k^{(2)}$, and for $r=k+1, \cdots, n$; $k=m$, $g_{r,s}=\psi_s^{(2)}$, $s=1, \cdots, k^{(2)}$, $g_{r,k^{(2)}+t}=\phi_t^{(2)}$, $t=1, \cdots, m-k^{(2)}$. Thus the $f_r\otimes g_{r,s}$ are linearly independent and such that either $f_r\epsilon\mathfrak{M}^{(1)\perp}$, or $g_{r,s}\epsilon\mathfrak{M}^{(2)\perp}$. Since there are $k^{(1)}(m-k^{(2)})+(n-k^{(1)})m=mn-k^{(1)}k^{(2)}$ of them and the dimensionality of $\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)}$ is $k^{(1)}k^{(2)}$ by Lemma 3.2.1, the $f_r\otimes g_{r,s}$ determine $(\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)})^\perp$.

3.3. Let $T^{(1)}$ be a linear transformation in $l_{p,n}$ and $T^{(2)}$ a linear transformation in $l_{p,m}$. Let $\phi_1^{(1)}, \cdots, \phi_n^{(1)}$ be n linearly independent elements of $l_{p,n}$ and $\phi_1^{(2)}, \cdots, \phi_m^{(2)}$, m linearly independent elements of $l_{p,m}$. Then the elements $\phi_i^{(1)}\otimes\phi_j^{(2)}$, $i=1, \cdots, n$, $j=1, \cdots, m$ determine $l_{p,mn}$ by Lemma 3.1.4 and are linearly independent. Hence given any set of mn elements $f_{i,j}\epsilon l_{p,mn}$, there exists a unique linear transformation T' such that

$$T'(\phi_i^{(1)}\otimes\phi_j^{(2)})=f_{i,j}\qquad (i=1,\cdots,n, j=1,\cdots,m).$$

Now let $f_{i,j}=T^{(1)}\phi_i^{(1)}\otimes T^{(2)}\phi_j^{(2)}$ and denote the corresponding T' by $T^{(1)}\otimes T^{(2)}$. Apparently this definition of $T^{(1)}\otimes T^{(2)}$ depends on the choice of the $\phi_i^{(1)}$'s and $\phi_j^{(2)}$'s, but this is not the case as is shown by the following

LEMMA 3.3.1. *If $f\epsilon l_{p,n}$, $g\epsilon l_{p,m}$, then $T^{(1)}\otimes T^{(2)}f\otimes g=T^{(1)}f\otimes T^{(2)}g$. Thus $T^{(1)}\otimes T^{(2)}$ does not depend on the choice of the $\{\phi_i^{(1)}\}$ or the $\{\phi_j^{(2)}\}$.*

The proof follows immediately from the definition of $T^{(1)}\otimes T^{(2)}$ and Lemma 3.1.1, (iv).

We also have

LEMMA 3.3.2. *A linear transformation T' of $l_{p,mn}$, equals $T^{(1)}\otimes T^{(2)}$ if and only if $T'f\otimes g=T^{(1)}f\otimes T^{(2)}g$ for every $f\epsilon l_{p,n}$, and $g\epsilon l_{p,m}$.*

The sufficiency of the condition follows from the definition. Lemma 3.3.1 implies its necessity.

LEMMA 3.3.3. $(T^{(1)}\otimes T^{(2)})^*=(T^{(1)})^*\otimes(T^{(2)})^*$

By Lemma 3.3.2, it suffices to show that if $f\epsilon l_{p',n}$, $g\epsilon l_{p',m}$, then $(T^{(1)}\otimes T^{(2)})^*f\otimes g=T^{(1)*}f\otimes T^{(2)*}g$. Now if $h\epsilon l_{p,mn}$, it follows from Lemma 3.1.4, that $h=\sum_{i=1}^n\sum_{j=1}^m a_{i,j}\phi_i^{(1)}\otimes\phi_j^{(2)}$ and by the definition of $T^{(1)}\otimes T^{(2)}$, $T^{(1)}\otimes T^{(2)}h=\sum_{i=1}^n\sum_{j=1}^m a_{i,j}T^{(1)}\phi_i^{(1)}\otimes T^{(2)}\phi_j^{(2)}$.

By the definition of T^* (cf. §2.3) and Lemma 3.1.1, (ii), we have

$$\begin{aligned}(T^{(1)}\otimes T^{(2)}h, f\otimes g) &= \sum_{i=1}^n\sum_{j=1}^m a_{i,j}(T^{(1)}\phi_i^{(1)}\otimes T^{(2)}\phi_j^{(2)}, f\otimes g)\\ &= \sum_{i=1}^n\sum_{j=1}^m a_{i,j}(T^{(1)}\phi_i^{(1)}, f)(T^{(2)}\phi_j^{(2)}, g)\\ &= \sum_{i=1}^n\sum_{j=1}^m a_{i,j}(\phi_i^{(1)}, T^{(1)*}f)(\phi_j^{(2)}, T^{(2)*}g)\\ &= \sum_{i=1}^n\sum_{j=1}^m a_{i,j}(\phi_i^{(1)}\otimes\phi_j^{(2)}, T^{(1)*}f\otimes T^{(2)*}g)\\ &= \left(\sum_{i=1}^n\sum_{j=1}^m a_{i,j}\phi_i^{(1)}\otimes\phi_j^{(2)}, T^{(1)*}f\otimes T^{(2)*}g\right).\\ &= (h, T^{(1)*}f\otimes T^{(2)*}g).\end{aligned}$$

Or for every h of $l_{p,mn}$,

$$(T^{(1)}\otimes T^{(2)}h, f\otimes g) = (h, T^{(1)*}f\otimes T^{(2)*}g).$$

The definition of T^*, then implies

$$(T^{(1)}\otimes T^{(2)})^*f\otimes g = T^{(1)*}f\otimes T^{(2)*}g$$

which is the desired result.

3.4. We have the following lemma.

LEMMA 3.4.1. (i) *If $E^{(1)}$ is a projection on $\mathfrak{M}^{(1)}\subseteq l_{p,n}$, and $E^{(2)}$ a projection on $\mathfrak{M}^{(2)}\subseteq l_{p,m}$, then $E^{(1)}\otimes E^{(2)}$ is a projection of $l_{p,mn}$ on $\mathfrak{M}^{(1)}\otimes\mathfrak{M}^{(2)}$.*

(ii) *Let $\phi_i^{(1)}$ and $\psi_i^{(1)}$ ($\phi_i^{(2)}$ and $\psi_i^{(2)}$) be in the same relation to $E^{(1)}$ ($E^{(2)}$) as ϕ_i and ψ_i are to E in Lemma 2.2.1, i.e.,*

$$E^{(1)}f = \sum_{i=1}^{k^{(1)}}(\psi_i^{(1)}, f)\phi_i^{(1)}; \qquad E^{(2)}f = \sum_{j=1}^{k^{(2)}}(\psi_j^{(2)}, f)\phi_j^{(2)}.$$

If E is the transformation on $l_{p,mn}$ defined by the equation

$$Eh = \sum_{i=1}^{k^{(1)}} \sum_{j=1}^{k^{(2)}} (\psi_i^{(1)} \otimes \psi_j^{(2)}, h)\phi_i^{(1)} \otimes \phi_j^{(2)},$$

then E is a projection of $l_{p,mn}$ on $\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$ and $E = E^{(1)} \otimes E^{(2)}$.

(iii) $E^* = E^{(1)*}E^{(2)*}$.

Since (ii) implies (i) and (ii) and Lemma 3.3.3 implies (iii), we need only prove (ii).

We have $(\psi_i^{(1)} \otimes \psi_j^{(2)}, \phi_l^{(1)} \otimes \phi_k^{(2)}) = (\psi_i^{(1)}, \phi_l^{(1)})(\psi_j^{(2)}, \phi_k^{(2)}) = \delta_{i,l}\delta_{j,k}$, by Lemma 3.1.1 and Lemma 2.2.1. By Lemma 3.2.1 the $\phi_l^{(1)} \otimes \phi_k^{(2)}$ determine $\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$. Lemma 2.2.1 now implies that E is a projection on $\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$.

It remains to show that $E = E^{(1)} \otimes E^{(2)}$. If $f \epsilon l_{p,n}$, $g \epsilon l_{p,m}$, then by Lemma 3.1.1, (ii) and (iv), and Lemma 2.2.1,

$$\begin{aligned} Ef \otimes g &= \sum_{i=1}^{k^{(1)}} \sum_{j=1}^{k^{(2)}} (\psi_i^{(1)} \otimes \psi_j^{(2)}, f \otimes g)\phi_i^{(1)} \otimes \phi_j^{(2)} \\ &= \sum_{i=1}^{k^{(1)}} \sum_{j=1}^{k^{(2)}} (\psi_i^{(1)}, f)(\psi_j^{(2)}, g)\phi_i^{(1)} \otimes \phi_j^{(2)} \\ &= \left(\sum_{i=1}^{k^{(1)}} (\psi_i^{(1)}, f)\phi_i^{(1)} \right) \otimes \left(\sum_{j=1}^{k^{(2)}} (\psi_j^{(2)}, g)\phi_j^{(2)} \right) = E^{(1)}f \otimes E^{(2)}g. \end{aligned}$$

Lemma 3.3.2 now implies that $E = E^{(1)} \otimes E^{(2)}$.

3.5. Next we prove

Lemma 3.5.1. *If $\mathfrak{M}^{(1)}$ in $l_{p,n}$ and $\mathfrak{M}^{(2)}$ in $l_{p,m}$ are in Situation* A (cf. §2.4), *then $\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$ is in Situation* A *with*

(a) $\phi_{(s-1)k^{(1)}+t} = \phi_t^{(1)} \otimes \phi_s^{(2)}, \quad t = 1, \cdots, k^{(1)}, \quad s = 1, \cdots, k^{(2)},$

$\psi_{(s-1)k^{(1)}+t} = \psi_t^{(1)} \otimes \psi_s^{(2)}, \quad t = 1, \cdots, k^{(1)}, \quad s = 1, \cdots, k^{(2)};$

(b) $h_{(s-1)r^{(1)}+t} = h_t^{(1)} \otimes h_s^{(2)}, \quad t = 1, \cdots, r^{(1)}, \quad s = 1, \cdots, r^{(2)};$

(c) $C = C^{(1)}C^{(2)};$

(d) $c_{(s-1)r^{(1)}+t} = c_t^{(1)} c_s^{(2)}, \quad t = 1, \cdots, r^{(1)}; \quad s = 1, \cdots, r^{(2)}.$

That the $\phi_t^{(1)} \otimes \phi_s^{(2)}$, $t = 1, \cdots, k^{(1)}$, $s = 1, \cdots, k^{(2)}$ are linearly independent and determine $\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$ has been shown in Lemma 3.1.3 and Lemma 3.2.1. The remaining statements of (a) were shown in the proof of Lemma 3.4.1.

To prove (b) we have $h_t^{(1)} \otimes h_s^{(2)} \epsilon \mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$ by definition (cf., §3.2). Also by Lemma 3.1.1 $\|h_t^{(1)} \otimes h_s^{(2)}\| = \|h_t^{(1)}\| \cdot \|h_s^{(2)}\| = 1$.

Now consider (c). If $i = (s-1)r^{(1)} + t$, then by Lemma 3.1.1, (ii); Lemma 3.4.1, (iii); Lemma 3.3.2; Lemma 3.1.1, (i); and §2.4, (c),

$$\begin{aligned}\|E^*\{h_i\}^{p-1}\| &= \|E^*\{h_t^{(1)} \otimes h_s^{(2)}\}^{p-1}\| = \|E^*\{h_t^{(1)}\}^{p-1} \otimes \{h_s^{(2)}\}^{p-1}\| \\ &= \|E^{(1)*} \otimes E^{(2)*}\{h_t^{(1)}\}^{p-1} \otimes \{h_s^{(2)}\}^{p-1}\| \\ &= \|E^{(1)*}\{h_t^{(1)}\}^{p-1} \otimes E^{(2)*}\{h_s^{(2)}\}^{p-1}\| \\ &= \|E^{(1)*}\{h_t^{(1)}\}^{p-1}\| \cdot \|E^{(2)*}\{h_t^{(2)}\}^{p-2}\|.\end{aligned}$$

We now prove (d). If $h \epsilon (\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)})^{\perp}$, then by Lemma 3.2.2, h is a linear combination of elements in the form $f^{(1)} \otimes g^{(2)}$, where either $f^{(1)}$ is in $\mathfrak{M}^{(1)\perp}$ or $g^{(2)}$ is in $\mathfrak{M}^{(2)\perp}$. Hence since

$$\sum_{i=1}^{r^{(1)}r^{(2)}} c_i(\{h_i\}^{p-1}, f)(h, \{E^*\{h_i\}^{p-1}\}^{p'-1}),$$

$f \epsilon \mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)}$, $h \epsilon (\mathfrak{M}^{(1)} \otimes \mathfrak{M}^{(2)})^{\perp}$, is linear in h, it is enough to show (d) for h in the form $f^{(1)} \otimes g^{(2)}$, where either $f^{(1)}$ is in $\mathfrak{M}^{(1)\perp}$ or $g^{(2)}$ is in $\mathfrak{M}^{(2)\perp}$. It is also linear in f; hence by §3.2 it suffices to show (d) for f in the form $\phi^{(1)} \otimes \phi^{(2)}$, $\phi_1 \epsilon \mathfrak{M}^{(1)}$, $\phi_2 \epsilon \mathfrak{M}^{(2)}$.

Now it was shown in the proof of (c) above that $E^*\{h_i\}^{p-1} = E^{(1)*}\{h_t^{(1)}\}^{p-1} \otimes E^{(2)*}\{h_s^{(2)}\}^{p-1}$. Then by Lemma 3.1.1, (iii),

$$\{E^*\{h_i\}^{p-1}\}^{p'-1} = \{E^{(1)*}\{h_t^{(1)}\}^{p-1}\}^{p'-1} \otimes \{E^{(2)*}\{h_s^{(2)}\}^{p-1}\}^{p'-1}$$

and $\{h_i\}^{p-1} = \{h_t^{(1)}\}^{p-1} \otimes \{h_s^{(2)}\}^{p-1}$. Hence by Lemma 3.1.1, (ii), we see that

$$\begin{aligned}&\sum_{i=1}^{r^{(1)}r^{(2)}} c_i(\{h_i\}^{p-1}, \phi^{(1)} \otimes \phi^{(2)})(f^{(1)} \otimes g^{(2)}, \{E^*\{h_i\}^{p-1}\}^{p'-1}) \\ &\quad = \left(\sum_{t=1}^{r^{(1)}} c_t^{(1)}(\{h_t^{(1)}\}^{p-1}, \phi^{(1)})(f^{(1)}, \{E^{(1)*}\{h_t^{(1)}\}^{p-1}\}^{p'-1}\right) \\ &\quad \times \left(\sum_{s=1}^{r^{(2)}} c_s^{(2)}(\{h_s^{(2)}\}^{p-1}, \phi^{(2)})(g^{(2)}, \{E^{(2)*}\{h_t^{(2)}\}^{p-1}\}^{p'-1}\right).\end{aligned}$$

Since either $f^{(1)} \epsilon \mathfrak{M}^{(1)\perp}$, or $g^{(2)} \epsilon \mathfrak{M}^{(2)\perp}$, this is zero for $\mathfrak{M}^{(1)}$ and $\mathfrak{M}^{(2)}$ are in Situation A.

3.6. Now it follows from Lemma 2.4.1 and Lemma 3.5.1, that we can show that $\lim_{n\to\infty} (\overline{C}(l_{p,n})) = \infty$ if we can find a $\mathfrak{M}$ in $l_{p,n}$ in Situation A (cf. §2.4, (c)). For let N be any integer >0. Then using Lemma 3.5.1 we can find a manifold $\mathfrak{N}_N$ in l_{p,n^N} in Situation A with $C_{\mathfrak{N}_N} = C_{\mathfrak{M}}^N$. Lemma 2.4.1 now implies that $C(\mathfrak{N}_N) \geqq C_{\mathfrak{M}}^N$ and since

$$\overline{C}(l_{p,n^N}) \geqq C(\mathfrak{N}_N), \quad \lim_{n\to\infty} (\overline{C}(l_{p,n})) = \infty.$$

As we remarked in §1.3, this implies that the answer to both (a) and (b) is negative.

Let $\mathfrak{M}$ be the manifold in $l_{p,3}$, determined by the vectors (1, 1, 0) and (0, 1, 1). Let

$$\phi_1 = (2^{-1/p}, 2^{-1/p}, 0), \qquad \phi_2 = (0, 2^{-1/p}, 2^{-1/p}),$$
$$\psi_1 = (2^{1/p+1}/3,\ 2^{1/p}/3,\ -\ 2^{1/p}/3), \quad \psi_2 = (-\ 2^{1/p}/3,\ 2^{1/p}/3,\ 2^{1/p+1}/3);$$

if $\alpha = 1/(2+2^p)^{1/p}$, $h_1 = (\alpha, -\alpha, -2\alpha)$, $h_2 = (\alpha, 2\alpha, \alpha)$, $h_3 = (2\alpha, \alpha, -\alpha)$. Also

$$C = ((2^{p'-1}+1)/3)^{1/p'}((2^{p-1}+1)/3)^{1/p},$$
$$c_1 = c_2 = c_3 = 1.$$

We show that $\mathfrak{M}$ is in Situation A (§2.4) and thus complete the proof.

We have (a) $(\phi_i, \psi_j) = \delta_{i,j}$.

(b) $h_1 = \alpha 2^{1/p}(\phi_1 - 2\phi_2)$, $h_2 = \alpha 2^{1/p}(\phi_1 + \phi_2)$, $h_3 = \alpha 2^{1/p}(2\phi_1 - \phi_2)$, and thus h_i is $\epsilon\mathfrak{M}$, $i=1, 2, 3$. We also have $\|h_i\| = 1$, $i = 1, 2, 3$.

Before showing (c) and (d) we make certain calculations. From the definitions in §2.1, we get $\{h_1\}^{p-1} = \alpha^{p-1}(1, -1, -2^{p-1})$, $\{h_2\}^{p-1} = \alpha^{p-1}(1, 2^{p-1}, 1)$, $\{h_3\}^{p-1} = (2^{p-1}, 1, 1)$, $(\{h_1\}^{p-1}, \phi_1) = (\{h_3\}^{p-1}, \phi_2) = 0$, $-(\{h_1\}^{p-1}, \phi_2) = (\{h_2\}^{p-1}, \phi_1) = (\{h_2\}^{p-1}, \phi_2) = (\{h_3\}^{p-1}, \phi_1) = 2^{-1}(1 + 2^{p-1})^{1/p}$.

By Lemma 2.3.1

$$E^*\{h_1\}^{p-1} = (\{h_1\}^{p-1}, \phi_1)\psi_1 + (\{h_1\}^{p-1}, \phi_2)\psi_2 = (\{h_1\}^{p-1}, \phi_2)\psi_2$$
$$= -\ 3^{-1}\,2^{-1/p'}(2^{p-1}+1)^{1/p}(-\ 1, 1, 2),$$
$$E^*\{h_2\}^{p-1} = (\{h_2\}^{p-1}, \phi_1)\psi_1 + (\{h_2\}^{p-1}, \phi_2)\psi_2 = 2^{-1}(2^{p-1}+1)^{1/p}(\psi_1 + \psi_2)$$
$$= 3^{-1}2^{-1/p'}(2^{p-1}+1)^{1/p}(1, 2, 1).$$

Similarly

$$E^*\{h_3\}^{p-1} = 3^{-1}2^{-1/p'}(2^{p-1}+1)^{1/p}(2, 1, -\ 1).$$

Finally (cf. §2.1)

$$\{E^*\{h_1\}^{p-1}\}^{p'-1} = -\ 3^{-(p'-1)}2^{-(p'-1)/p'}(2^{p-1}+1)^{(p'-1)/p}(-\ 1, 1, 2^{p'-1})$$
$$\equiv -\ K(-\ 1, 1, 2^{p'-1})$$
$$\{E^*\{h_2\}^{p-1}\}^{p'-1} = K(1, 2^{p'-1}, 1)$$
$$\{E^*\{h_3\}^{p-1}\}^{p'-1} = K(2^{p'-1}, 1, -\ 1).$$

(c) By direct calculation, we obtain $\|E^*\{h_1\}^{p-1}\| = \|E^*\{h_2\}^{p-1}\| = \|E^*\{h_3\}^{p-1}\| = C$ using the above. For $p \neq 2$, C is >1 since by the Hölder inequality,

$$6C = (2^p+2)^{1/p}(2^{p'}+2)^{1/p'} \geqq 2\cdot 2 + 2^{1/p}2^{1/p'} = 6,$$

where the equality sign holds only for $p=2$.

(d) Now if $f\epsilon\mathfrak{M}^\perp$, then $f = k\tilde{\varphi}$, where $\tilde{\phi} = (1, -1, 1)$. Thus

$$(\{E^*\{h_1\}^{p-1}\}^{p'-1}, f) = Kk(2 - 2^{p'-1}),$$
$$(\{E^*\{h_2\}^{p-1}\}^{p'-1}, f) = Kk(2 - 2^{p'-1}),$$
$$(\{E^*\{h_3\}^{p-1}\}^{p'-1}, f) = Kk(2^{p'-1} - 2).$$

We can now verify by a direct calculation that (d) holds.

Conclusion. Our results permit us to conclude that

There exists a manifold $\mathfrak{M}_0$ in l_p and L_p such that there exists no biorthogonal set $\{\phi_i, \psi_i\}$ where $\{\phi_i\}$ is a basis for $\mathfrak{M}_0$ (cf. (B), Chapter VII, p. 110, §3), *while the expansion*

$$(*) \qquad \sum_{i=1}^{\infty} a_i\phi_i, \qquad a_i = (f, \psi_i),$$

converges for each $f \epsilon l_p$ or L_p.

Let us suppose that (*) converges for every f. Let $\mathfrak{M}$ be the manifold determined by the ϕ_i's. The ϕ_i's are a basis for $\mathfrak{M}$ (cf. (B), loc. cit.) for if $f \epsilon \mathfrak{M}$, then

$$f = \sum_{i=1}^{\infty} a_i\phi_i, \qquad a_i = (f, \psi_i)$$

by (B), Chapter VII, Theorem 2, p. 107.

We will show that under these circumstances $C(\mathfrak{M})$ is $<\infty$. For let E be the transformation defined by the equation

$$Ef = \sum_{i=1}^{\infty} (f, \psi_i)\phi_i.$$

Ef is defined for every f since we assume that the series is convergent for every f. The same assumption implies that E is limited for the partial sums are uniformly limited (cf. (B), Chapter VII, Theorem 2 and Theorem 5). E is obviously additive and homogeneous. If $f \epsilon \mathfrak{M}$, $Ef = f$ by the above and the range of E is included in $\mathfrak{M}$, $E\Lambda = \mathfrak{M}$ and $E^2 = E$. Thus E is a projection of Λ on $\mathfrak{M}$. Hence $C(\mathfrak{M})$ is $<\infty$.

Our construction also permits us to show that no statement concerning the norms of the ϕ_i and ψ_i will insure convergence by itself. We can assume that $\|\phi_i\| = 1$ for every i. The least possible value for $\|\psi_i\|$ is then 1 since $(\phi_i, \psi_i) = 1$ and $|(\phi_i, \psi_i)| \leqq \|\phi_i\| \cdot \|\psi_i\|$. We will show that *there exists in both l_p and L_p a biorthogonal set $\{\phi_i, \psi_i\}$ for which $\|\phi_i\| = \|\psi_i\| = 1$ and for which the associated expansion* (*) *does not always converge.*

It is a consequence of the proof of Lemma 1.3.1 that if such a set exists in l_p, there must be a similar one in L_p. So we need only consider l_p. Owing to the nature of biorthogonal sets in l_p, we need only consider the case $1 < p < 2$.

Let $\mathfrak{M}$ be the manifold of §3.6, and let $f_1=(2^{-1/p}, 2^{-1/p}, 0)=\phi_1, f_2=(\alpha, -\alpha, -2\alpha)=h_1$. We have $(f_1, \{f_2\}^{p-1})=(\{f_1\}^{p-1}, f_2)=0$; $(f_1, \{f_1\}^{p-1})=(f_2, \{f_2\}^{p-1})=1$, $\|f_1\|=\|f_2\|=\|\{f_1\}^{p-1}\|=\|\{f_2\}^{p-1}\|=1$. Of course f_1 and $f_2 \epsilon l_{p,3}$, $\{f_1\}^{p-1}$ and $\{f_2\}^{p-1}\epsilon l_{p',3}$.

We define $f_{i_1}\otimes f_{i_2}\otimes \cdots \otimes f_{i_n}$, $i_j=1, 2$; $j=1, \cdots, n$, as an element of $l_{p,3^n}$ as follows: $f_{i_1}\otimes f_{i_2}$ in $l_{p,3^2}$ has already been defined (cf. §3.1). Let us suppose that $f_{i_1}\otimes \cdots \otimes f_{i_{n-1}}$ in $l_{p,3^{n-1}}$ has been defined. We define $f_{i_1}\otimes f_{i_2}\otimes \cdots \otimes f_{i_{n-1}}\otimes f_{i_n}$ as $(f_{i_1}\otimes \cdots \otimes f_{i_{n-1}})\otimes f_{i_n}$ in $l_{p,3^n}$ using §3.1. Let $\mathfrak{M}_1=\mathfrak{M}$, $\mathfrak{M}_n$ in $l_{p,3^n}$ be $\mathfrak{M}_{n-1}\otimes\mathfrak{M}$. Then by successive applications of Lemma 3.2.1, we see that the set of elements $f_{i_1}\otimes \cdots \otimes f_{i_n}$ determine $\mathfrak{M}_n$.

By Lemma 1.1.2, $l_p=\sum_{\alpha=1}^{\infty}\oplus l_{p,3^\alpha}$. Let $\mathfrak{P}$ be the closed linear manifold in $\sum_{\alpha=1}^{\infty}\oplus l_{p,3^\alpha}$ consisting of those elements $\{g_1, g_2, \cdots\}$ for which g_α is $\epsilon\mathfrak{M}_\alpha$ for every α. Let S consist of those elements which are such that every $g_\alpha=0$ except for one g_n and $g_n=f_{i_1}\otimes \cdots \otimes f_{i_n}$. Let S' consist of elements of $l_{p'}$ in the form $\{g\}^{p-1}$, $g\epsilon S$. Since as we have seen above the $f_{i_1}\otimes \cdots \otimes f_{i_n}$ determine $\mathfrak{M}_n$, S determines $\mathfrak{P}$.

Now the sets S and S' are denumerable and it is easily seen by using Lemma 3.1.1 that with a suitable enumeration they form a biorthogonal set with $\|\phi_i\|=\|\psi_i\|=1$. Since S determines $\mathfrak{P}$, we see from the above that this series (*) cannot converge always if $C(\mathfrak{P})=\infty$.

Let C be as in §3.6, (c). By repeated applications of Lemma 3.5.1 and then using Lemma 2.4.1, one may prove that $C(\mathfrak{M}_n)\geqq C^n$. It follows from the proof of Lemma 1.2.3 that $C(\mathfrak{P})$ is $\geqq C(\mathfrak{M}_n)\geqq C^n$ for every n and since C is >1, $p\neq 2$ this implies that $C(\mathfrak{P})=\infty$. As we have remarked above, this proves our statement.

Incidently we have explicitly constructed a manifold $\mathfrak{P}$ in l_p, for which there exists no projection. Lemma 1.3.1 indicates how we can find a $\mathfrak{P}$ in L_p with the corresponding property.

In L_p, the space of complex-valued functions whose pth power is summable, the situation is the same. As pointed out in a previous paper by the writer,† the theorems given in (B) and used here can be generalized to the complex case. Chapter 1 of this paper also falls into this category. Some variations are necessary in Chapters 2 and 3 but they are not basic.

Finally it should be pointed out that the negative answer to (a) and (b) precludes the possibility of a spectral theory in l_p and L_p similar to the theory of self-adjoint operators in Hilbert space.

† These Transactions, vol. 39 (1936), pp. 83–100.

Columbia University,
New York, N. Y.

A CORRECTION†

BY

I. M. SHEFFER

I have discovered an error (and some minor misstatements) in my paper *A local solution of the difference equation* $\Delta y(x) = F(x)$ *and of related equations* (these Transactions, vol. 39 (1936), pp. 345–379). It is my purpose here to rectify that error. The vulnerable point occurs from relation (7.1) to Theorem 7.1. In order to secure convergence for series (7.2), it was asserted that the Mittag-Leffler theorem could be applied to (7.6), giving a meromorphic function $Z(x; x^*)$. Now the "poles" of this function are at the points

$$x = x^* + \omega_1 + \sum_{j=2}^{k} n_j(\omega_1 - \omega_j),$$

where $n_2, \cdots, n_k$ range independently from $-\infty$ to $+\infty$. But there is no reason to suppose (as I did) that these "poles" have no limit points in the *finite* plane.‡ If there *are* finite limit points, then we cannot conclude that $L[Z]$ is an entire function, and therefore we cannot apply Theorem 6.6 (Carmichael's theorem). Consequently, Theorem 7.1 will not follow from the argument given.

It becomes necessary to rewrite §7. Fortunately it is possible to give a rigorous treatment which is simpler than the old. We now indicate in what way the old paper is to be revised.

(1) In Theorem 4.4, in place of " $\cdots$ *and in this circle* $y(x)$ $\cdots$ " read " $\cdots$ *and in the lens-region* $y(x)$ $\cdots$."

(2) In Theorem 4.6, in place of " $\cdots$ *which in this circle satisfies* $\cdots$ " read " $\cdots$ *which in a lens-region about* $x = \alpha$ *satisfies* $\cdots$."

(3) §5 should have as heading: "5. **A geometric lemma.**"

(4) Omit all of §5 beginning with the following line (just preceding Lemma 5.2: "We turn now to two lemmas $\cdots$." (This portion is now unnecessary.)

(5) In §6 omit the two paragraphs shortly after (6.17), beginning with "To treat the general case, in which $F(x)$ $\cdots$," and ending with "This implies no loss of generality of equation (6.3)." (This portion is now unnecessary.)

(6) §7 is to be replaced by the following:

† Received by the editors, June 27, 1936.

‡ For certain values of $n_2, \cdots, n_k$, the coefficient $b_{n_2, \cdots, n_k}$ is zero, so that the corresponding "pole" does not actually occur. Conceivably, enough "poles" may be absent so that there are no finite limit points, but there seems to be no way of enumerating the "poles" that are present.

7. **The general case.** To treat the general case where $F(x)$ is merely analytic, we modify the method of §4. Instead of directly finding a solution of the homogeneous equation $L[y]=0$, we shall obtain more than one merom rphic solution of equation

$$L[y] = \frac{1}{x-\alpha}. \tag{7.1}$$

Let C be the unique circle, center at x^* and radius ρ^*, assured us by Lemma 5.1. As a consequence of uniqueness there are seen to be precisely two possibilities: Either

Case I. There are at least two points P_j on C, and of these at least one pair, say P_1, P_2, are *diametrically opposite.* Or,

Case II. There are at least three points P_j on C; of these no two are diametrically opposite, but at least one triad of them (say P_1, P_2, P_3) forms an *acute-angled triangle.*

In Case I draw circles of radius σ (any number greater than ρ^*) with centers at P_1, P_2. These circles will form a lens-region enclosing x^*, and as $\sigma\rightarrow\rho^*$ the lens-region (including boundary) will close down on x^* as a unique limit point. That is, given any neighborhood $\mathfrak{N}$ of x^*, there exists a σ such that the lens-region of radius σ will lie with its boundary wholly interior to $\mathfrak{N}$.

Now consider Case II. A simple geometric argument will show that of P_1, P_2, P_3, there will be one, say P_1, such that: (i) P_2 and P_3 are on opposite sides of the diameter d_1 through P_1; and (ii) if d_1' is the diameter perpendicular to d_1, then P_2 and P_3 are on that side of d_1' opposite P_1. Now let σ be any number greater than ρ^*, and draw circles of radius σ and centers P_1, P_2, P_3. These circles form a curvilinear triangle and another simple geometric argument will show that as $\sigma\rightarrow\rho^*$ this triangle closes down on the single point x^*. Hence again, if $\mathfrak{N}$ is any neighborhood of x^*, there is a σ such that the curvilinear triangle of radius σ lies with its boundary wholly inside $\mathfrak{N}$.

As a solution of (7.1) assume the series

$$Y_1(x;\alpha) = \sum_{n_2,\cdots,n_k=0}^{+\infty} \frac{b_{n_2,\cdots,n_k}}{\left[x-(\alpha+\omega_1)-\sum_{j=2}^{k} n_j(\omega_1-\omega_j)\right]}. \tag{7.2}$$

On substituting into (7.1) we get

$$\frac{1}{x-\alpha} = \sum_{n_2,\cdots,n_k=0}^{+\infty} \frac{\alpha_1 b_{n_2\cdots n_k} + \sum_{j=2}^{k}\alpha_j b_{n_2,\cdots,n_{j-1},n_j-1,n_{j+1},\cdots,n_k}}{\left[x-\alpha-\sum_{j=2}^{k} n_j(\omega_1-\omega_j)\right]},$$

where all b's with a negative subscript are zero. This condition is formally fulfilled if we choose the b's to satisfy

$$\alpha_1 b_{n_2\cdots n_k} + \sum_{j=2}^{k} \alpha_j b_{n_2,\cdots,n_{j-1},n_j-1,n_{j+1},\cdots,n_k} \tag{7.3}$$
$$= \begin{cases} 1, & \text{for } n_2 = \cdots = n_k = 0; \\ 0, & \text{for } n_2, \cdots, n_k = 0, 1, 2, \cdots \text{ (but not all zero)}. \end{cases}$$

We find from $(n_2, \cdots, n_k) = (0, \cdots, 0)$ that $b_{0,\cdots,0} = 1/a_1$. Then on choosing $n_2 + \cdots + n_k = 1$ in all possible ways we find that the $b_{n_2\cdots n_k}$'s $(n_2 + \cdots + n_k = 1)$ are uniquely determined; then the $b_{n_2\cdots n_k}$'s for $n_2 + \cdots + n_k = 2$; etc. That is, there is a *unique set* of b's for which relations (7.3) hold. These values we choose for the coefficients in (7.2).

(7.2) is a *formal* series; there is no reason to suppose that it converges. But if we examine its formal "poles," namely, the points

$$x = \alpha + \omega_1 + \sum_{j=2}^{k} n_j(\omega_1 - \omega_j),$$

we find that they have no limit point in the finite plane.† The classic theorem of Mittag-Leffler is therefore applicable. Set $u = x - \alpha$, so that

$$Y_1(x; \alpha) = \sum_{n_2,\cdots,n_k=0}^{+\infty} \frac{b_{n_2\cdots n_k}}{\left[u - \omega_1 - \sum_{j=2}^{k} n_j(\omega_1 - \omega_j)\right]}.$$

Then there exist polynomials $P_{n_2\cdots n_k}(u)$ such that

$$Z_1(u) = \sum_{n_2,\cdots,n_k=0}^{+\infty} \left(\frac{b_{n_2\cdots n_k}}{\left[u - \omega_1 - \sum_{j=2}^{k} n_j(\omega_1 - \omega_j)\right]} + P_{n_2\cdots n_k}(u) \right) \tag{7.4}$$

defines a meromorphic function whose only poles are at

$$u = \omega_1 + \sum_{j=2}^{k} n_j(\omega_1 - \omega_j)$$

(with corresponding residues $b_{n_2\cdots n_k}$), the series converging uniformly and

† For consider the point P_1 (i.e., $-\omega_1$), which lies on C. If L_1 is the tangent line to C at P_1, then $P_2, \cdots, P_k$ all lie on the same side of L_1. Hence if we start at P_1 and lay off vectors $\sum_{j=2}^{k} n_j(\omega_1 - \omega_j)$, we see (since $n_j \geqq 0$) that the ends of the vectors all lie on this same side of L_1 (save for $n_2 = \cdots = n_k = 0$). If we take components of these vectors in the direction perpendicular to the line L_1, we see that the ends of the vectors go off to infinity as any n_j becomes infinite, so that no finite limit point is possible.

absolutely in every bounded region (the poles in this region being deleted).

On applying the operator L term-wise to the series for Z_1 (as is permissible) we obtain

$$L[Z_1(x-\alpha)] = \sum_{n_2\cdots n_k=0}^{+\infty}\left(\frac{A_{n_2\cdots n_k}}{\left[x-\alpha-\sum_{j=2}^{k} n_j(\omega_1-\omega_j)\right]} + L[P_{n_2\cdots n_k}(x-\alpha)]\right)$$

where $A_{n_2\cdots n_k}$ is the left member of (7.3), so that

$$L[Z_1(x-\alpha)] = \frac{1}{x-\alpha} + \sum_{n_2\cdots n_k=0}^{+\infty} L[P_{n_2\cdots n_k}(x-\alpha)]. \tag{7.5}$$

The right-hand side converges for all x (the point $x=\alpha$ is singular only for the term $1/(x-\alpha)$). That is, the series on the right is an entire function. Now by Theorem 6.6 the equation

$$L[g_1(u)] = \sum_{n_2\cdots n_k=0}^{+\infty} L[P_{n_2\cdots n_k}(u)] \tag{7.6}$$

has an entire function solution $g_1(u)$. Consequently

$$W_1(u) = Z_1(u) - g_1(u) \tag{7.7}$$

is a meromorphic function satisfying the equation

$$L[W_1(x-\alpha)] = \frac{1}{x-\alpha}; \tag{7.8}$$

its only poles are simple poles at the points

$$x = \alpha + \omega_1 + \sum_{j=2}^{k} n_j(\omega_1-\omega_j) \qquad (n_2, \cdots, n_k = 0, 1, 2, \cdots),$$

and the corresponding residues are $b_{n_2\cdots n_k}$ *as given by* (7.3).

In the above work concerning Y_1, Z_1, and W_1, the point P_1 (i.e., the number $-\omega_1$) was preferred over all the others. But in Case I point P_2 is also on C, and in Case II points P_2, P_3 are on C; and these points may equally well be used as was P_1. We thus get, according to the case, one or two unique formal series†

$$Y_2(x;\alpha) = \sum_{n_1,n_3,\cdots,n_k=0}^{+\infty} \frac{c_{n_1n_3\cdots n_k}}{\left[x-(\alpha+\omega_2)-\sum_{j=1,3,\cdots,k} n_j(\omega_2-\omega_j)\right]},$$

† The c's and d's satisfy recurrence relations similar to (7.3).

$$Y_3(x;\alpha) = \sum_{n_1,n_2,n_4,\cdots,n_k=0}^{+\infty} \frac{d_{n_1n_2n_4\cdots n_k}}{\left[x-(\alpha+\omega_3)-\sum_{j=1,2,4,\cdots,k} n_j(\omega_3-\omega_j)\right]};$$

one or two functions $Z_2(x-\alpha)$, $Z_3(x-\alpha)$; and one or two meromorphic functions $W_2(x-\alpha)$, $W_3(x-\alpha)$ satisfying (respectively)

$$L[W_2(x-\alpha)] = \frac{1}{x-\alpha}, \qquad L[W_3(x-\alpha)] = \frac{1}{x-\alpha},$$

with respective simple poles at

$$x = \alpha+\omega_2+\sum_{j=1,3,\cdots,k} n_j(\omega_2-\omega_j), \qquad x = \alpha+\omega_3+\sum_{j=1,2,4,\cdots,k} n_j(\omega_3-\omega_j),$$

and residues $c_{n_1n_3\cdots n_k}$, $d_{n_1n_2n_4\cdots n_k}$.

Case I. Here P_1, P_2 are diametrically opposite on C. Let σ be only slightly larger than ρ^*, and with radius σ and centers P_1, P_2, draw a lens-region $\mathfrak{L}$ around x^*. Let the bounding arcs of $\mathfrak{L}$ be C_1, C_2 (C_i being that arc with center at P_i). Let α remain on C_1. Since the point $x^*+\omega_1$ is where the point P_2 would be if C were translated so that x^* falls at the origin, it is seen that as α traverses C_1, $\alpha+\omega_1$ will trace a small arc (in the neighborhood of $x^*+\omega_1$) of a circle of radius σ and center the origin. Hence from our knowledge of the position of the poles

$$x = \alpha+\omega_1+\sum_{j=2}^{k} n_j(\omega_1-\omega_j),$$

we can say, if σ is sufficiently close to ρ^*, that $W_1(x-\alpha)$ is analytic about the origin in a circle of radius exceeding ρ^*. That is,

$$W_1(x-\alpha) = \sum_{n=0}^{\infty} B_{1n}(\alpha)x^n, \tag{7.9}$$

where the $B_{1n}(\alpha)$ are analytic functions in the neighborhood of $\alpha=x^*$ (and in particular for α on C_1), and where there is a number $\sigma_1>\rho^*$, *independent of α on C_1*, such that (7.9) converges uniformly for α on C_1 and x in $|x|\leqq\sigma_1$.

An analogous statement applies to $W_2(x-\alpha)$ for α on C_2:

$$W_2(x-\alpha) = \sum_{n=0}^{\infty} B_{2n}(\alpha)x^n, \tag{7.10}$$

uniformly convergent for α on C_2 and x in $|x|\leqq\sigma_2$, where σ_2 is some number exceeding ρ^*. (The $B_{2n}(\alpha)$ are analytic functions in a neighborhood of x^* containing C_2.)

Case II. P_1, P_2, P_3 are on C. Again choosing σ only slightly larger than ρ^*, we obtain a curvilinear triangle $\mathfrak{L}$ of radius σ by drawing arcs C_1, C_2, C_3 with centers P_1, P_2, P_3. For α on C_i the argument used in Case I applies, giving us (7.9), (7.10) or

$$W_3(x-\alpha) = \sum_{n=0}^{\infty} B_{3n}(\alpha)x^n, \tag{7.11}$$

uniformly convergent for α on C_3 and x in $|x| \leqq \sigma_3$, where σ_3 is some number greater than ρ^*. (The $B_{3n}(\alpha)$ are analytic in a neighborhood of x^* containing C_3.)

For any value of i ($i=1$ or 2 in Case I and 1, 2, or 3 in Case II), $\limsup |B_{in}(\alpha)|^{1/n} \leqq 1/\sigma_i \leqq 1/\sigma < 1/\rho^*$, where σ = smallest of σ_1, σ_2, σ_3 and α is on C_i. It follows from Theorem 6.2 and Corollary 6.1 that the series $\sum_{n=0}^{\infty} B_{in}(\alpha)A_n(x)$ converges uniformly for x in some curvilinear polygon $\mathfrak{L}$ about x^* and α on C_i, where $\mathfrak{L}$ can be chosen independent of i. But this series is what we get when we apply L term-wise to the $W_i(x-\alpha)$ series. From this follows

THEOREM 7.1. *According to the case, if a lens-region $\mathfrak{L}$ or a curvilinear triangle $\mathfrak{L}$ be drawn† about x^*, with radius σ sufficiently near to ρ^*, then*

$$\frac{1}{x-\alpha} = \sum_{n=0}^{\infty} B_{jn}(\alpha)A_n(x), \qquad \begin{cases} j = 1, 2 & \textit{in Case I}, \\ j = 1, 2, 3 & \textit{in Case II}, \end{cases} \tag{7.11}$$

the convergence being uniform in x and α for α on C_j and x in some neighborhood $\mathfrak{N}$ of x^*. ($\mathfrak{N}$ can be chosen independent of j.)

Now let $F(x)$ be analytic about $x=x^*$. Then there exists a lens or triangle $\mathfrak{L}$ around x^* lying (together with its boundary) wholly in the region of analyticity of $F(x)$, and with radius so close to ρ^* that Theorem 7.1 applies. Let x be in the region $\mathfrak{N}$ of Theorem 7.1. Multiply (7.11) by $F(\alpha)$ and integrate over $\mathfrak{L}$. This gives

$$F(x) = \sum_{n=0}^{\infty} f_n A_n(x), \tag{7.12}$$

where

$$f_n = -\sum_{\substack{j=1,2 \\ \text{or } j=1,2,3}} \frac{1}{2\pi i}\int_{C_j} F(\alpha)B_{jn}(\alpha)d\alpha; \tag{7.13}$$

and (7.12) converges uniformly in $\mathfrak{N}$. We thus have

† The points P_j will of course be the centers of the arcs forming $\mathfrak{L}$.

THEOREM 7.2. *If $F(x)$ is analytic about $x=x^*$, it has a convergent A_n-expansion, given by* (7.12).

Combining this with Theorem 6.1:

THEOREM 7.3. *A necessary and sufficient condition that a function $F(x)$ have an (convergent) A_n-expansion is that it be analytic at $x=x^*$.*

By Theorem 6.2, $\limsup|f_n|^{1/n}<1/\rho^*$, so that the series

$$(7.14)\qquad y(x) = \sum_0^\infty f_n x^n$$

converges in $|x| <\rho^*+\epsilon$, for some $\epsilon>0$. On applying L to (7.17) we get

$$L[y(x)] = \sum f_n L[x^n] = \sum f_n A_n(x) = F(x),$$

so that we have

THEOREM 7.4. *If $F(x)$ is analytic about $x=x^*$, then the function $y(x)$ of* (7.14) *is analytic in a circle (about $x=0$) of radius greater than ρ^*, and for all x in a sufficiently small curvilinear polygon (about $x=x^*$) $y(x)$ satisfies the equation*

$$(6.3)\qquad L[y(x)] = F(x).$$

The point x^* is of course significant for A_n-expansions, but not for equation (6.3). For let $F(x)$ be analytic about $x=c$, and define $G(x)=F(x+c-x^*)$. $G(x)$ is analytic about $x=x^*$ and therefore there exists a function $z(x)$, analytic at x^*, such that $L[z(x)]=G(x)$. Consequently, the function $y(x)=z(x-c+x^*)$ satisfies $L[y(x)]=F(x)$, and we have the final

THEOREM 7.5. *If $F(x)$ is analytic about $x=c$, there exists a function $y(x)$, analytic about $x=c-x^*$ in a circle of radius exceeding ρ^*, such that for all x in a sufficiently small neighborhood (curvilinear polygon) of $x=c$, $y(x)$ satisfies equation* (6.3).

PENNSYLVANIA STATE COLLEGE,
STATE COLLEGE, PA.

ON SOME FUNCTIONALS, II*

BY

S. SAKS

1. This article is primarily intended to correct the mistakes in §§4 and 5 of the former paper by the author *On some functionals*.† On the whole these errors do not affect the theorems themselves with the exception of the obviously false remarks that Theorem 4 and a part of Theorem 3 of S. F. hold for the space R of characteristic functions. In the present note we complete the gaps in the proofs of those theorems‡ and slightly strengthen Theorem 3 of S. F. in the part concerning the equal continuity of the operations considered (cf. below Theorem 3, (ii)). We also prove two theorems which were not stated in S. F., namely, Theorems 1 and 4 so as to obtain a more symmetric set of results.

2. We shall recall briefly the notation. Let $\{\xi_n(t)\}$ be a sequence of measurable functions on a measurable set U. Then

(i) $\{\xi_n(t)\}$ is *bounded in measure* on U if to every $\eta>0$ there corresponds a number $M=M(\eta)$ such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon U;\ |\xi_n(t)| > M] < \eta \qquad (n = 1, 2, \cdots);$$

(ii) $\{\xi_n(t)\}$ *converges in measure* on U if to every η there corresponds a $k=k(\eta)$ such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon U;\ |\xi_n(t) - \xi_m(t)| > \eta] < \eta$$

whenever $n>k$, $m>k$;

(iii) $\{\xi_n(t)\}$ *has property* $B(\epsilon)$ on U, if there exists a number M independent of n, such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon U;\ |\xi_n(t)| > M] < \epsilon \qquad (n = 1, 2, \cdots);$$

(iv) $\{\xi_n(t)\}$ *has property* $C(\epsilon)$ on U,§ if to every $\eta>0$ there corresponds a $k=k(\eta)$ such that

* Presented to the Society, September 5, 1936; received by the editors April 27, 1936.

† These Transactions, vol. 35 (1932), pp. 549–556. This paper will be referred to as S. F.

‡ For the proof of Parts (i) and (ii) of Theorem 3 of S. F. see also the recent book by Kaczmarz and Steinhaus, *Theorie der Orthogonalreihen*, Monografje Matematyczne, Warszawa-Lwow, 1935, pp. 24–26. The author's attention was called to these mistakes by Kaczmarz and by Gowurin. (Cf. Gowurin, *On sequences of indefinite integrals*, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 930–936.)

§ We shall write $\{\xi_n(t)\}\varepsilon B(\epsilon)$ or $\varepsilon C(\epsilon)$ to indicate that the sequence $\{\xi_n(t)\}$ has property $B(\epsilon)$ or $C(\epsilon)$ respectively.

$$\text{meas}\, \underset{t}{E}[t\varepsilon U;\ |\xi_n(t) - \xi_m(t)| > \eta] < \epsilon, \qquad n > k,\ m > k.$$

In what follows we denote by I a measurable set of finite measure (e.g., an interval) and by E a Banach space. We shall consider sequences of functions $\{\xi_n(x, t)\}$ depending on $x\varepsilon E$ and $t\varepsilon I$. For each x fixed in E, $\xi_n(x, t)$ are finite and measurable functions of t in I. On the other hand, as functions of x, they are supposed to be linear operations on E; i.e., additive and continuous. This means that $\xi_n(x_1+x_2, t)=\xi_n(x_1, t)+\xi_n(x_2, t)$ almost everywhere on I, whenever $x_1\varepsilon E$, $x_2\varepsilon E$ and that $\xi_n(x, t)$ tends in measure to 0 on I when $|x|\to 0$ (for n fixed). In considering $\xi_n(x, t)$ as operations on E we shall often write $\xi_n(x, t)\equiv F_n(x)$.

The operations $F_n(x)=\xi_n(x, t)$ will be said to be *equally continuous* if to every $\eta>0$ there corresponds an $r>0$ such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon I;\ |\xi_n(x, t)| > \eta] < \eta \qquad (n = 1, 2, \cdots),$$

whenever $|x| < r$. They will be said to be *equally continuous with respect to a set* $U \subset I$ if the condition above is satisfied with I replaced by U.

Finally $\{T_k\}$ $(k=1, 2, \cdots)$ will denote a sequence of measurable sets in I such that the characteristic functions of the sets T_k form an everywhere dense set in the space of all characteristic functions defined on I. Thus for each set Q and every $\eta>0$ there is a set T_k, $k=k(Q, \eta)$, such that meas $(Q-QT_k)<\eta$ and meas $(T_k-QT_k)<\eta$.

3. We have

LEMMA 1. *Let* $\sum_1^\infty \epsilon_k$ *be a converging series of positive numbers,* $\{P_q\}$ *a sequence of measurable sets in* I *and*

$$P = \limsup_k P_k = \prod_{k=1}^{\infty} \sum_{q=k}^{\infty} P_q. \tag{3.1}$$

Then, if a sequence $\{\xi_n(t)\}\varepsilon B(\epsilon_k)$ *on* P_k $(k=1, 2, \cdots)$, *it is bounded in measure on* P; *and if a sequence* $\{\xi_n(t)\}\varepsilon C(\epsilon_k)$ *on* P_k $(k=1, 2, \cdots)$, *it converges in measure on* P.

Let

$$Q_k = \sum_{q=k}^{\infty} P_q, \qquad Q_{k,s} = \sum_{q=k}^{s} P_q,$$

and let η be an arbitrary positive number. Choose k_0 and next s_0, so that

$$\sum_{i=k_0}^{\infty} \epsilon_i < \frac{\eta}{2}, \qquad \text{meas}\,(Q_{k_0} - Q_{k_0,s_0}) < \frac{\eta}{2}. \tag{3.2}$$

Suppose now that the sequence under consideration $\varepsilon B(\epsilon_k)$ on P_k. Then there is a number M such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon Q_{k_0,s_0};\ |\ \xi_n(t)\ |\ > M] < \sum_{q=k_0}^{s_0} \epsilon_q < \frac{\eta}{2} \qquad (n = 1, 2, \cdots);$$

thus, by (3.1) and (3.2)

$$\text{meas}\, \underset{t}{E}[t\varepsilon P;\ |\ \xi_n(t)\ |\ > M] \leqq \text{meas}\, \underset{t}{E}[t\varepsilon Q_{k_0};\ \ \xi_n(t)\ |\ > M] < \eta$$
$$(n = 1, 2, \cdots);$$

i.e., the sequence $\{\xi_n(t)\}$ is bounded in measure on P.

Next suppose that the sequence $\{\xi_n(t)\}\varepsilon C(\epsilon_k)$ on P_k. Then there is a positive integer n_0 such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon Q_{k_0,s_0};\ |\ \xi_n(t) - \xi_m(t)\ |\ > \eta] < \sum_{i=k_0}^{s_0} \epsilon_i < \frac{\eta}{2}, \qquad n \geqq n_0, \quad m \geqq n_0,$$

and so, by (3.1) and (3.2)

$$\text{meas}\, \underset{t}{E}[t\varepsilon P;\ |\ \xi_n(t) - \xi_m(t)\ |\ > \eta] \leqq \text{meas}\, \underset{t}{E}[t\varepsilon Q_{k_0};\ |\ \xi_n(t) - \xi_m(t)\ |\ > \eta] < \eta,$$
$$n \geqq n_0,\ m \geqq n_0.$$

LEMMA 2. *If a sequence $\{\xi_n(x, t)\}\varepsilon B(\epsilon)$ on a measurable set $P \subset I$ for every x belonging to a set H of the second category in E, then, to every $\eta > 0$, there corresponds an $r > 0$ such that for an arbitrary $x\varepsilon E$ with $|x| < r$,*

$$\text{(3.3)} \qquad \text{meas}\, \underset{t}{E}[t\varepsilon P;\ |\ \xi_n(x, t)\ |\ > \eta] < 3\epsilon \qquad (n = 1, 2, \cdots).$$

Let H_m denote the set of all $x\varepsilon H$ such that

$$\text{meas}\, \underset{t}{E}[t\varepsilon P;\ |\ \xi_n(x, t)\ |\ > m] < \epsilon \qquad (n = 1, 2, \cdots).$$

We have $H = \sum_1^\infty H_m$, and so there is a positive integer m_0 such that H_{m_0} is of the second category. Since $\xi_n(x, t)$ is continuous in x on E, the inequalities

$$\text{meas}\, \underset{t}{E}[t\varepsilon P;\ |\ \xi_n(x, t)\ |\ > m_0] \leqq \epsilon \qquad (n = 1, 2, \cdots)$$

hold for all $x\varepsilon \overline{H}_{m_0}$, and consequently for all x in a sphere $K_0 \subset \overline{H}_{m_0}$. Let r_0 be the radius of K_0. By the linearity of the operations $\xi_n(x, t)$, we have

$$\text{meas}\, \underset{t}{E}\, [t\varepsilon P;\ |\ \xi_n(x, t)\ |\ > 2m_0] \leqq 2\epsilon \qquad (n = 1, 2, \cdots),$$

whenever $|x| < r_0$, and therefore condition (3.3) is satisfied with $r = \eta r_0/(2m_0)$.

4. We now prove the following theorem:

THEOREM 1. *For any given sequence $\{\xi_n(x, t)\}$ there exists a set A in I such that*

(i) *the sequence $\{\xi_n(x, t)\}$ is bounded in measure on A for all $x\varepsilon E$;*

(ii) *the operations* $F_n(x)=\xi_n(x, t)$ *are equally continuous with respect to* A;

(iii) *for every* x *in* E, *except perhaps for a set of the first category in* E, *the sequence* $\{\xi_n(x, t)\}$ *is bounded in measure on no set in* $I-A$ *of positive measure.*

Denote by α_0 the least upper bound of all numbers α with the property that there exists a set $H(\alpha)$ of the second category in E such that for each x in $H(\alpha)$ the sequence $\{\xi_n(x, t)\}$ is bounded in measure on a set of measure $>\alpha$. Let q be a positive integer. We shall prove first that there exists a set H_q of the second category in E and a set P_q in I of measure $>\alpha_0-1/q^2$ such that

$$\{\xi_n(x, t)\}\,\varepsilon B(1/q^2) \text{ on } P_q, \quad x\varepsilon H_q. \tag{4.1}$$

Indeed to every $x\varepsilon H(\alpha_0-1/q^2)$ we can attach a set T_k, $k=k(x)$ (see §2), such that

$$\{\xi_n(x, t)\}\,\varepsilon B(1/q^2) \text{ on } T_k, \quad \text{meas } T_k > \alpha_0 - \frac{1}{q^2}. \tag{4.2}$$

Let $H_{q,k}$ denote the set of all $x\varepsilon H(\alpha_0-1/q^2)$ for which (4.2) is satisfied for a fixed k. Since $H(\alpha_0-1/q^2)=\sum_{k=1}^{\infty}H_{q,k}$ is of the second category, there exists a $k=k_q$ such that H_{q,k_q} is of the second category. On putting $H_q=H_{q,k_q}$ and $P_q=T_{k_q}$, we see at once that H_q and P_q satisfy condition (4.1).

We shall prove next that for every x in E

$$\{\xi_n(x, t)\}\,\varepsilon B(8/q^2) \text{ on } P_q. \tag{4.3}$$

To show this consider an arbitrary point $x_0\varepsilon \overline{H}_q$. By (4.1) and Lemma 2

$$\text{meas } \underset{t}{E}[t\varepsilon P_q;\ |\xi_n(x_0 - x_1, t)| > 1] < \frac{3}{q^2} \qquad (n = 1, 2, \cdots), \tag{4.4}$$

if $|x_0-x_1|$ is sufficiently small. Hence we can choose $x_1\varepsilon H_q$ so as to satisfy (4.4). Since $\{\xi_n(x_1, t)\}\,\varepsilon B(1/q^2)$ on P_q this implies that $\{\xi_n(x_0, t)\}\,\varepsilon B(4/q^2)$ on P_q. But x_0 is an arbitrary element of the set $\overline{H}_q$ which certainly contains interior points (a sphere), whence it follows that $\{\xi_n(x, t)\}\,\varepsilon B(8/q^2)$ on P_q for all x in E.

Now set $A=\prod_{k=1}^{\infty}\sum_{q=k}^{\infty}P_q$. It results from (4.3) and Lemma 1 that the sequence $\{\xi_n(x, t)\}$ is bounded in measure on A, i.e., that A satisfies condition (i) of the theorem.

Next let η be an arbitrary positive number, and let, as in the proof of Lemma 1,

$$Q_k = \sum_{q=k}^{\infty} P_q, \qquad Q_{k,s} = \sum_{q=k}^{s} P_q. \tag{4.5}$$

Let k_0 and s_0 be positive integers such that

$$(4.6) \qquad 24/(k_0 - 1) < \frac{\eta}{2}, \quad \text{meas}\,(Q_{k_0} - Q_{k_0, s_0}) < \frac{\eta}{2}.$$

By (4.3), Lemma 2 and (4.6), there exists an $r = r(\eta)$ such that, for $|x| < r$,

$$\text{meas}\, \underset{t}{E}[t \in Q_{k_0, s_0};\ |\xi_n(x, t)| > \eta] < 24 \sum_{q=k_0}^{s_0} q^{-2} < \frac{\eta}{2} \qquad (n = 1, 2, \cdots),$$

and thus, again by (4.6)

$$\text{meas}\, \underset{t}{E}[t \in A;\ |\xi_n(x, t)| > \eta] \leqq \text{meas}\, \underset{t}{E}[t \in Q_{k_0};\ |\xi_n(x, t)| > \eta] < \eta$$

which is condition (ii).

Finally, in order to prove condition (iii) suppose that there is a set H of the second category in E such that for every $x \in H$ the sequence $\{\xi_n(x, t)\}$ is bounded in measure on a set $Q(x) \subset I - A$ of positive measure. Then, for every $x \in H$ the sequence $\{\xi_n(x, t)\}$ would be bounded in measure on $A + Q(x)$. Now since

$$\text{meas}\,[A + Q(x)] > \text{meas}\, A \geqq \limsup_q\ \text{meas}\, P_q \geqq \lim_q\ (\alpha_0 - 1/q^2) = \alpha_0,$$

we have

$$H = \sum_{n=1}^{\infty} H_n, \quad H_n = \underset{x}{E}[x \in H; \text{meas}\,\{A + Q(x)\} \geqq \alpha_0 + 1/n].$$

Since at least one of the sets H_n is of the second category, this contradicts the definition of α_0.

5. Next we prove

THEOREM 2. *There exists a set B in I such that*

(i) *for all x in E, the sequence $\{\xi_n(x, t)\}$ converges in measure on B;*

(ii) *for every x in E, except perhaps for a set of the first category in E, the sequence $\{\xi_n(x, t)\}$ converges in measure on no set in $I - B$ of positive measure.*

First observe that for every $x \in E$, except perhaps for a set of the first category, the sequence $\{\xi_n(x, t)\}$ does not converge in measure on any set of positive measure contained in $I - A$ (where A is the set defined in Theorem 1). Hence to prove Theorem 2 we may assume that $I = A$, or, which amounts to the same (see Theorem 1, (ii)) that the operations $F_n(x) = \xi_n(x, t)$ are equally continuous with respect to the whole set I.

From now on we shall follow the line of the proof of Theorem 1. Denote by β_0 the least upper bound of all numbers β with the property that there exists a set $H(\beta)$ of the second category in E such that for every $x \in H(\beta)$ the sequence $\{\xi_n(x, t)\}$ converges in measure on a set of measure $> \beta$. Let q be an arbitrary

positive integer, and let $H_{q,k}$ denote the set of all $x\epsilon H(\beta_0-1/q^2)$ such that $\{\xi_n(x,t)\}\epsilon C(1/q^2)$ on a set T_k (see §2) of measure $>\beta_0-1/q^2$. Clearly $H(\beta_0-1/q^2)=\sum_{k=1}^{\infty}H_{q,k}$, and thus there is a $k=k_q$ such that H_{q,k_q} is of the second category. Hence upon putting $H_q=H_{q,k_q}$ and $P_q=T_{k_q}$ we see that there exists a set H_q of the second category in E and a set P_q in I of measure $>\beta_0-1/q^2$ such that

$$\{\xi_n(x,t)\}\epsilon C(1/q^2) \text{ on } P_q,\quad x\epsilon H_q. \tag{5.1}$$

Now as in the proof of Theorem 1 we shall show that for every x in E

$$\{\xi_n(x,t)\}\epsilon C(6/q^2) \text{ on } P_q. \tag{5.2}$$

Indeed, let $x_0\epsilon \overline{H}_q$ and η be an arbitrary positive number. The equal continuity of the operations $F_n(x)$ implies the existence of an $x_1\epsilon H_q$ sufficiently near to x_0 such that

$$\text{meas}\ \underset{t}{E}\left[t\epsilon I;\ |\xi_n(x_1-x_0,t)|>\frac{\eta}{6}\right]<\frac{1}{q^2}\qquad (n=1,2,\cdots). \tag{5.3}$$

Again, in view of (5.1) there exists a positive integer n_0 such that, for $n\geqq n_0$, $m\geqq n_0$,

$$\text{meas}\ \underset{t}{E}\left[t\epsilon P_q;\ |\xi_m(x_1,t)-\xi_n(x_1,t)|>\frac{\eta}{6}\right]<\frac{1}{q^2}.$$

Thus, by (5.3),

$$\text{meas}\ \underset{t}{E}\left[t\epsilon P_q;\ |\xi_m(x_0,t)-\xi_n(x_0,t)|>\frac{\eta}{2}\right]<\frac{3}{q^2},\qquad m\geqq n_0,\ n\geqq n_0. \tag{5.4}$$

Let K_0 be an arbitrary sphere contained in $\overline{H}_q$ and r_0 the radius of K_0. Since (5.4) holds for all elements $x_0\epsilon K_0$, we have, by the linearity of the transformations $F_n(x)$,

$$\text{meas}\ \underset{t}{E}[t\epsilon P_q;\ |\xi_n(x,t)-\xi_m(x,t)|>\eta]<\frac{6}{q^2},\qquad m\geqq n_0,\ n\geqq n_0,\ |x|<r_0.$$

This means however that $\{\xi_n(x,t)\}\epsilon C(6/q^2)$ on P_q whenever $|x|<r_0$ and consequently for every x in E.

Now let

$$B=\prod_{k=1}^{\infty}\sum_{q=k}^{\infty}P_q.$$

By Lemma 1 the sequence $\{\xi_n(x,t)\}$ converges in measure on B for every x. Hence, the set B satisfies condition (i). To establish condition (ii) suppose

that there is a set H of the second category such that for every $x\varepsilon H$ the sequence $\{\xi_n(x, t)\}$ converges in measure on a set $Q(x) \subset I-B$ of positive measure. Then, for every $x\varepsilon H$ the sequence would converge in measure on $B+Q(x)$. But

$$\text{meas } [B + Q(x)] > \text{meas } B \geqq \limsup_q \text{ meas } P_q \geqq \beta_0,$$

which contradicts the definition of the number β_0 (cf. the proof of Theorem 1). Condition (ii) is thus established.

6. We now have

THEOREM 3. *There exists a set C in I such that*

(i) *for all x in E,* $\sup_n |\xi_n(x, t)| < \infty$ *almost everywhere in* C;

(ii) *to every* $\eta>0$ *there corresponds an* $r>0$ *such that*

$$\text{meas } \underset{t}{E}[t\varepsilon C;\ \sup_n \ |\ \xi_n(x, t)\ | > \eta] < \eta, \qquad |\ x\ | < r;$$

(iii) *for every x in E, except perhaps for a set of the first category in E,*

$$\sup_n |\ \xi_n(x, t)\ | = \infty$$

almost everywhere in $I-C$.

Let γ_0 be the least upper bound of all numbers γ such that the set $H(\gamma)$ of x for which

$$\text{meas } \underset{t}{E}[\sup_n \ |\ \xi_n(x, t)\ | < \infty] > \gamma$$

is of the second category. Let q be an arbitrary positive integer and let $H_{q,p,k}$ denote the set of all $x\varepsilon H(\gamma_0-1/q^2)$ such that

$$\text{meas } \underset{t}{E}[t\varepsilon T_k;\ \sup_n \ |\ \xi_n(x, t)\ | > p] < \frac{1}{q^2}, \qquad \text{meas } T_k > \gamma_0 - \frac{1}{q^2}.$$

It is clear that $H(\gamma_0-1/q^2)=\sum_{p,k=1}^{\infty} H_{q,p,k}$ and so there exist integers $k=k_q$ and $p=p_q$ such that H_{q,p_q,k_q} is of the second category. Put $H_q=H_{q,p_q,k_q}$ and $P_q=T_{k_q}$. Thus for every $x\varepsilon H_q$

$$\text{meas } \underset{t}{E}[t\varepsilon P_q;\ \sup_n \ |\ \xi_n(x, t)\ | > p_q] < \frac{1}{q^2}$$

while meas $P_q>\gamma_0-1/q^2$. Hence, by continuity of $\xi_n(x, t)$, for every $x\varepsilon \overline{H}_q$ and all s,

$$\text{meas } \underset{t}{E}[t\varepsilon P_q;\ \sup_{n\leqq s} \ |\ \xi_n(x, t)\ | > p_q] \leqq \frac{1}{q^2},$$

and therefore, for all $x \varepsilon \overline{H}_q$,

$$(6.1) \qquad \operatorname{meas} \underset{t}{E}[t \varepsilon P_q;\ \sup_n |\xi_n(x, t)| > p_q] \leqq \frac{1}{q^2}.$$

Let K_0 be a sphere contained in $\overline{H}_q$. If r_0 is the radius of K_0, (6.1) implies in view of the linearity of $\xi_n(x, t)$,

$$\operatorname{meas} \underset{t}{E}[t \varepsilon P_q;\ \sup_n |\xi_n(x, t)| > 2p_q] \leqq \frac{2}{q^2},$$

no matter what $x \varepsilon E$, provided $|x| < r_0$, and consequently, for any $\sigma > 0$,

$$(6.2) \qquad \operatorname{meas} \underset{t}{E}[t \varepsilon P_q;\ \sup_n |\xi_n(x, t)| > \sigma] \leqq \frac{2}{q^2}, \qquad |x| < \frac{r_0 \sigma}{2p_q}.$$

Now, again by linearity of $\xi_n(x, t)$ it results from (6.2) that for every $x \varepsilon E$, $\sup_n |\xi_n(x, t)| < \infty$ for all $t \varepsilon P_q$, with the exception of at most a subset of measure $\leqq 2/q^2$. Hence, upon putting $C = \prod_{k=1}^{\infty} \sum_{q=k}^{\infty} P_q$ we see at once that, for every $x \varepsilon E$, $\sup_n |\xi_n(x, t)| < \infty$ almost everywhere in C. Thus C satisfies condition (i) of the theorem.

Next, as in the proof of Lemma 1, let

$$Q_k = \sum_{q=k}^{\infty} P_q, \qquad Q_{k,s} = \sum_{q=k}^{s} P_q.$$

Let η be an arbitrary positive number, and let k_0 and s_0 be positive integers such that

$$(6.3) \qquad \frac{2}{k_0 - 1} < \frac{\eta}{2}, \quad \operatorname{meas}\ (Q_{k_0} - Q_{k_0, s_0}) < \frac{\eta}{2}.$$

In virtue of (6.2) there is an $r > 0$ such that whenever $|x| < r$,

$$\operatorname{meas} \underset{t}{E}[t \varepsilon Q_{k_0, s_0};\ \sup_n |\xi_n(x, t)| > \eta] \leqq \sum_{q=k_0}^{s_0} 2q^{-2} < \frac{\eta}{2}$$

and therefore, by (6.3),

$$\operatorname{meas} \underset{t}{E}[t \varepsilon C;\ \sup_n |\xi_n(x, t)| > \eta]$$
$$\leqq \operatorname{meas} \underset{t}{E}[t \varepsilon Q_{k_0};\ \sup_n |\xi_n(x, t)| > \eta] < \eta, \qquad |x| < r.$$

Thus condition (ii) for C is established. Finally condition (iii) follows at once from the definition of γ_0, since

$$\operatorname{meas} C \geqq \limsup_q \operatorname{meas} P_q \geqq \gamma_0.$$

7. Next we have

THEOREM 4. *There exists a set D in I such that*

(i) *for all x in E the sequence* $\{\xi_n(x, t)\}$ *converges almost everywhere on D;*

(ii) *for every x in E except perhaps for a set of the first category, the sequence* $\{\xi_n(x, t)\}$ *diverges almost everywhere on* $I-D$.

First observe that for every x, except perhaps for a set of the first category, the sequence $\{\xi_n(x, t)\}$ diverges almost everywhere on $I-C$, where C is the set defined in Theorem 3. Thus, without loss of generality, we may assume that $I=C$.

For a fixed $x\epsilon E$ let $\Gamma(x)$ denote the subset of I on which $\{\xi_n(x, t)\}$ converges. Further let $H(\delta)$ denote the set of all $x\epsilon E$ such that meas $\Gamma(x)>\delta$, $\delta>0$, and let δ_0 be the upper bound of the numbers δ for which $H(\delta)$ is of the second category.

Now, let q be an arbitrary positive integer and $H_{q,k}$ the set of all $x\epsilon H(\delta_0-1/q^2)$ such that

$$\text{meas } T_k > \delta_0 - \frac{1}{q^2}, \quad \text{meas } (T_k - T_k\Gamma(x)) < \frac{1}{q^2}.$$

We thus have $H(\delta_0-1/q^2)=\sum_{k=1}^{\infty}H_{q,k}$, and so there exists a $k=k_q$ such that $H_{q,k}$ is of the second category. Put $H_q=H_{q,k_q}$ and $P_q=T_{k_q}$. Then

$$(7.1) \qquad \text{meas } (P_q - P_q\Gamma(x)) \leqq \frac{1}{q^2}, \qquad x\epsilon H_q,$$

but we shall show that the latter inequality holds for every $x\epsilon \overline{H}_q$. Indeed, let $x_0\epsilon\overline{H}_q$ and η be an arbitrary positive number. In virtue of Theorem 3 (condition (ii) with $C=I$) we can find an $x_1\epsilon H_q$ such that

$$\text{meas } \underset{t}{E}\Big[t\epsilon I;\ \sup_n \ |\xi_n(x_0 - x_1, t)| > \eta\Big] < \eta.$$

Again, by (7.1) there is a positive integer $n_0=n_0(\eta)$ such that

$$\text{meas } \underset{t}{E}\Big[t\epsilon P_q;\ \sup_{n\geqq n_0} \ |\xi_n(x_1, t) - \xi_{n_0}(x_1, t)| > \eta\Big] \leqq \frac{1}{q^2} + \eta,$$

whence

$$\text{meas } \underset{t}{E}\Big[t\epsilon P_q;\ \sup_{n\geqq n} \ |\xi_n(x_0, t) - \xi_{n_0}(x_0, t)| > 3\eta\Big] \leqq \frac{1}{q^2} + 3\eta.$$

Since $\eta>0$ is arbitrary this implies the convergence of the sequence $\{\xi_n(x, t)\}$ for every $x\epsilon\overline{H}_q$ and for all $t\epsilon P_q$, with the exception of at most a subset of

measure $\leqq 1/q^2$. Since $\overline{H}_q$ certainly contains a sphere the same holds for every $x\epsilon E$ by the linearity of $\xi_n(x, t)$. Therefore, for each $x\epsilon E$ the sequence converges almost everywhere on the set $D=\prod_{k=1}^{\infty}\sum_{q=k}^{\infty}P_q$. Thus the set D satisfies condition (i) of the theorem and condition (ii) follows at once since

$$\text{meas } D \geqq \limsup_q \text{ meas } P_q \geqq \delta_0.$$

8. In view of Theorems 1, 2, 3, and 4 to every sequence $\{\xi_n(x, t)\}$ of the type considered there are attached four sets A, B, C, and D. We obviously have $B \subset A$ and $D \subset C$ (except for sets of measure zero). However, on account of condition (ii) of Theorem 1 (or 3) we have $B=A$ (or $D=C$) whenever E contains an everywhere dense set E_1 such that the sequence $\{\xi_n(x, t)\}$ converges in measure (or converges almost everywhere) in I for every x belonging* to E_1. It follows that if E is a separable space then from the sequence $\{\xi_n(x, t)\}$ a subsequence $\{\xi_n'(x, t)\}$ may be extracted which converges almost everywhere in BC for every $x\epsilon E$; in fact it is sufficient to select the required subsequence so that it converges almost everywhere in B for every x belonging to an everywhere dense, denumerable set in E.

9. The theorems of the preceding sections do not hold† in general, if the linear space E is replaced by the space R of characteristic functions. However, for the latter we have the following theorem:

THEOREM 5. (i) *If for all x belonging to a set H of the second category in R the sequence $\{F_n(x)=\xi_n(x, t)\}$, $x\epsilon R$, $t\epsilon I$, converges in measure on I, then the operations $F_n(x)$ are equally continuous in I, i.e., for every $\eta>0$ there exists an $r>0$ such that*

$$\text{meas } \underset{t}{E}[t\epsilon I;\ |\xi_n(x, t)| > \eta] \leqq \eta \qquad (n = 1, 2, \cdots),$$

whenever $|x| < r$.

(ii) *If for all x belonging to a set H of the second category in R, the sequence $\{\xi_n(x, t)\}$ converges almost everywhere on I, then for every $\eta>0$ there exists an $r>0$ such that*

$$\text{meas } \underset{t}{E}[t\epsilon I;\ \sup_n |\xi_n(x, t)| > \eta] < \eta,$$

whenever $|x| < r$.

We shall confine ourselves to the proof of statement (ii) of the theorem. The proof of statement (i) is the same as that of the lemma of S. F., p. 555.

* Cf. Banach, *Sur la convergence presque partout de fonctionnelles linéaires*, Bulletin des Sciences Mathématiques, vol. 50 (1926), pp. 27–32, 36–43; Saks, *Sur les fonctionnelles de M. Banach et leur application aux développements des fonctions*, Fundamenta Mathematicae, vol. 10 (1927), pp. 186–196; Mazur and Orlicz, *Über Folgen linearer Operationen*, Studia Mathematica, vol. 4 (1933), pp. 152–157; esp. p. 157; Kaczmarz and Steinhaus, op. cit., pp. 177–178.

† See Gowurin, loc. cit.

Let η be any positive number and k a positive integer; let H_k denote the set of all $x\varepsilon H$ such that

$$\operatorname{meas} \underset{t}{E}\left[t\varepsilon I;\ \sup_{n\geqq k} \left| \xi_n(x,t) - \xi_k(x,t) \right| > \frac{\eta}{4}\right] \leqq \frac{\eta}{4}.$$

We have $H=\sum_1^\infty H_k$ and thus there exists a $k=k_0$ such that H_{k_0} is of the second category. Hence

$$\operatorname{meas} \underset{t}{E}\left[t\varepsilon I;\ \sup_{n\geqq k_0} \left| \xi_n(x,t) - \xi_{k_0}(x,t) \right| > \frac{\eta}{4}\right] \leqq \frac{\eta}{4},$$

for each $x\varepsilon H_{k_0}$, and therefore, by continuity (cf. the proof of Theorem 3), for all $x\varepsilon \overline{H}_{k_0}$. Let K_0 be a sphere which is contained in $\overline{H}_{k_0}$ and let r_0 be its radius. For every element $x\varepsilon R$, $|x|<r_0$, there are two elements x_1 and x_2 in R such* that $x_1\varepsilon K_0$, $x_2\varepsilon K_0$ and $x_1=x_2+x$. It is readily seen that, no matter what x is, $|x|<r_0$,

$$\operatorname{meas} \underset{t}{E}\left[t\varepsilon I;\ \sup_{n\geqq k_0} \left| \xi_n(x,t) - \xi_{k_0}(x,t) \right| > \frac{\eta}{2}\right] \leqq \frac{\eta}{2}. \tag{9.1}$$

Let now $r<r_0$ be a positive number such that

$$\operatorname{meas} \underset{t}{E}\left[t\varepsilon I;\ \left| \xi_n(x,t) \right| > \frac{\eta}{2}\right] \leqq \frac{\eta}{2k_0} \qquad (n=1,2,\cdots,k_0),$$

whenever $|x|<r$. Then in view of (9.1), for every x, $|x|<r$,

$$\operatorname{meas} E_t\left[t\varepsilon I;\ \sup_n \left| \xi_n(x,t) \right| > \eta\right] \leqq \eta.$$

* We put $x_1=x+x_0(1-x)$, $x_2=x_0(1-x)$ where x_0 is the center of the sphere K_0.

The University of Warsaw,
Warsaw, Poland

FUNCTIONS DEFINED BY SEQUENCES OF INTEGRALS AND THE INVERSION OF APPROXIMATE DERIVED NUMBERS*

BY

R. L. JEFFERY

1. Introduction. Let the function $f(x)$ be measurable on (a, b), and let $s=s_1(x), s_2(x), \cdots$ be a sequence of summable functions tending to $f(x)$ almost everywhere. If $f(x)$ is summable and x a point of (a, b), conditions under which

$$(1) \qquad \lim_{n\to\infty}\int_a^x s_n(x)dx = \int_a^x f(x)dx$$

have been determined.† It is easy to construct sequences for which the limit on the left side of (1) exists and is different from the right side. In the present paper necessary and sufficient conditions are obtained for the existence of this limit in terms of the sequence $s_n(x)$. It turns out that the limit function $F(x)$ is independent of $f(x)$. In fact if $F(x)$ is an arbitrary continuous function, there exists a sequence of summable functions $s_n(x)$ tending to zero everywhere, for which $\int_a^x s_n(x)dx$ tends to $F(x)$ everywhere. If $F_1(x), F_2(x), \cdots$ is a sequence of measurable functions, there is a sequence $s_n(x)$ tending to zero everywhere for which almost everywhere the set of limits of $\int_a^x s_n(x)dx$ is the sequence $F_1(x), F_2(x), \cdots$. If $\int_e s_n(x)dx$ is bounded in n and e, e any measurable subset of (a, b), then $F(x)$, when it exists, is of bounded variation on (a, b). Conversely, if $F(x)$ is of bounded variation, there exists a function $f(x)$ and a sequence of summable functions $s_n(x)$ tending everywhere to $f(x)$ for which $\int_a^x s_n(x)dx$ tends to $F(x)$, and for which $\int_e s_n dx$ is bounded in n and e. This is of some interest for the reason that it provides a characterization of functions of bounded variation which can be extended to functions of any number of variables.

It is possible for the limit of the left side of (1) to exist when the function $f(x)$ is not summable. As an aid in the study of this situation we introduce the following conventions:

* Presented to the Society, December 27, 1934, and April 20, 1935; received by the editors May 3, 1935, and in revised form November 13, 1935, and July 11, 1936.

† W. H. Young, *Term by term integration of oscillating series*, Proceedings of the London Mathematical Society, ser. 2, vol. 8, pp. 99–116. Jeffery, *The integrability of a sequence of functions*, these Transactions, vol. 33 (1931), pp. 433–440. T. H. Hildebrandt, *On the interchange of limit and Lebesgue integral for a sequence of functions*, ibid., pp. 441–443. In the second of these papers further references are given to the literature on the subject.

Let $f(x)$ be measurable on (a, b). If there exists a sequence of summable functions $s_n(x)$ tending to $f(x)$ almost everywhere and a continuous function $F(x)$ such that $\int_a^x s_n(x)dx$ tends to $F(x)$, then $f(x)$ is *integrable in the sequence sense to* $F(x)$, $F(x)=S(f, a, x)$. If $s_n(x)$ can be so determined that $s_n(x)=f$ on E_n, $s_n(x)=0$ elsewhere, E_n contains E_{n-1}, and mE_n tends to $b-a$, then $f(x)$ is *totally integrable in the sequence sense to* $F(x)$, $F(x)=TS(f, a, x)$.

In the light of the foregoing statements concerning the limits of $\int_a^x s_n dx$ with s_n tending to zero, it is clear that $S(f, a, x)$ is not uniquely defined. If f is finite almost everywhere, and almost everywhere is the approximate derivative of the continuous function $F(x)$, then $S(f, a, x)$ exists for which $S(f, a, x)=F(x)-F(a)$. If $f(x)$ is summable, then $TS(f, a, x)$ is uniquely defined, and $TS(f, a, x)=\int_a^x f(x)dx$. If $f(x)$ is integrable in the generalized Denjoy sense, and $F(x)=\int_a^x f(x)dx$, then there exists $TS(f, a, x)=F(x)$, but in this case $TS(f, a, x)$ is not uniquely defined. It is probable that if $f(x)$ is integrable in the generalized Denjoy sense and $TS(f, a, x)$ can be determined which is (ACG),* then $TS(f, a, x)=F(x)$. Some information is given in regard to this point, but so far it has not been possible to obtain all the facts.

2. **The limit function of the sequence of integrals.** Let f be summable, and let s_n be a sequence of summable functions tending to f almost everywhere. For an arbitrary positive number η let $E(l, \eta)$ be the part of (a, x) for which $|f-s_n|<\eta$, $n \geqq l$, and let $C(l, \eta)$ be the complement of $E(l, \eta)$ on (a, x). These sets are measurable, and as l increases $mC(l, \eta)$ tends to zero. If for a fixed x

$$F(x) = \lim_{n\to\infty} \int_a^x s_n dx = \lim_{n\to\infty} \int_{E(l,\eta)} s_n dx + \lim_{n\to\infty} \int_{C(l,\eta)} s_n dx,$$

then for l sufficiently large the first limit on the right is arbitrarily near to $\int_a^x f dx$. Hence if $F(x)$ exists, we have

$$F(x) = \int_a^x f\,dx + \lim_{l\to\infty}\left[\lim_{n\to\infty} \int_{C(l,\eta)} s_n dx\right].$$

We thus get:

THEOREM I. *A necessary and sufficient condition for the existence of* $F(x)$ *is the existence of*

$$\lim_{l\to\infty} \int_{C(l,\eta)} s_n dx, \qquad n \geqq l.$$

---

* Generalized absolutely continuous. $F(x)$ is (ACG) on (a, b) if it is continuous, and (a, b) can be separated into a finite or denumerable set of sets $E_1, E_2, \cdots$ such that $F(x)$ is absolutely continuous on each E_n; Saks, *Théorie de l'Intégral*, Warsaw, 1933, p. 152, §9.

Let $U(n, \delta)$, $L(n, \delta)$ be respectively the least upper bound and greatest lower bound of $\int_e s_n$ for all e on (a, x) with $me < \delta$. The reasoning used by Hildebrandt* can be modified to give:

THEOREM II. *A necessary and sufficient condition that $F(x)$ exists is that*

$$\lim_{\delta \to 0} \lim_{n \to \infty} [U(n, \delta) + L(n, \delta)] = K(x).$$

Let $g = g_1, g_2, \cdots$ be any subsequence of s, and let $U(g, n, \delta)$, $L(g, n, \delta)$ be the least upper bound and greatest lower bound respectively of $\int_e g_i$, $me < \delta$, $i = 1, 2, \cdots, n$.

THEOREM III. *A necessary and sufficient condition that $F(x)$ exist is that*

$$\lim_{\delta \to 0} \lim_{n \to \infty} [U(g, n, \delta) + L(g, n, \delta)] = K,$$

where K depends on x but is independent of g.

The proof of this can be accomplished by the methods of our previous paper.† We give here a proof which is much simpler and which includes the results of that paper as a special case.

Suppose $F(x)$ exists. Then by Theorem I of the present paper we have

$$\lim_{l \to \infty} \int_{C(l,\eta)} s_n dx = K, \qquad n \geqq l. \tag{1}$$

Suppose there is some g of s such that for every $\delta' > 0$ there exists $\delta < \delta'$ and a sequence of positive integers $n_1, n_2, \cdots$ such that

$$U(g, n_i, \delta) + L(g, n_i, \delta) > K + \lambda, \qquad \lambda > 0. \tag{2}$$

Fix l so that

$$\left| \int_{C(l,\eta)} s_n dx - K \right| < \frac{\lambda}{4}, \qquad n \geqq l. \tag{3}$$

For a fixed n, $U(g, n, \delta)$ tends to zero as δ tends to zero. As a result of this, together with (2), it follows that there exists δ, n_i, n_i', and a set e, such that $l < n_i' \leqq n_i$, $me < \delta$, and

$$\int_e g_{n_i'} dx > U(g, n_i, \delta) - \frac{\lambda}{4}.$$

If C^+ and C^- are the parts of $C(l, \eta)$ for which $g_{n_i'} \geqq 0$, $g_{n_i'} < 0$, respectively, reasoning similar to that used by Hildebrandt‡ shows that

* Loc. cit., pp. 441–442.

† These Transactions, loc. cit.

‡ Loc. cit., p. 442, lines 1–7.

$$\int_{C^+} g_{n_i'} dx > U(g, n_i, \delta) - \frac{\lambda}{2},$$

while the definition of $L(g, n, \delta)$ gives

$$\int_{C^-} g_{n_i'} dx \geqq L(g, n_i, \delta).$$

Hence, since $C^+ + C^- = C(l, \eta)$,

$$\int_{C(l,\eta)} g_{n_i'} dx > U(g, n_i, \delta) + L(g, n_i, \delta) - \frac{\lambda}{2} > K + \frac{\lambda}{2}.$$

Since $n_i' > l$ this contradicts (3). Similar reasoning leads to a contradiction, if in (2) $K+\lambda$ is replaced by $K-\lambda$ and the inequality sign is reversed. This shows that the condition is necessary.

Next suppose that the condition holds and $F(x)$ does not exist. If there exists a subsequence $g = g_1, g_2, \cdots$ of s such that $\int_a^x g_n dx$ tends to $\pm\infty$, the method of the first part of the theorem can be used to show that

$$\lim_{\delta\to 0} \lim_{n\to\infty} [U(g, n, \delta) + L(g, n, \delta)] = \pm\infty.$$

If no subsequence exists and $F(x)$ does not exist, there then exists two subsequences $g = g_1, g_2, \cdots$ and $h = h_1, h_2, \cdots$ of s such that

$$\lim_{n\to\infty} \int_a^x g_n dx = G > \lim_{n\to\infty} \int_a^x h_n dx = H.$$

By Theorem I,

$$G = \int_a^x f\,dx + G', \qquad H = \int_a^x f\,dx + H'.$$

Hence $G' > H'$. By the first part of this theorem

$$\lim_{\delta\to 0} \lim_{n\to\infty} [U(g, n, \delta) + L(g, n, \delta)] = G',$$

$$\lim_{\delta\to 0} \lim_{n\to\infty} [U(h, n, \delta) + L(h, n, \delta)] = H'.$$

Since $G' > H'$ the hypotheses are contradicted, and the sufficiency of the condition follows.

If s_n is such that on (a, b), $\int_e s_n dx$ is bounded in n and e, then for a fixed x and g

$$\lim_{\delta\to 0} \lim_{n\to\infty} U(g, n, \delta) = U(g, x), \qquad \lim_{\delta\to 0} \lim_{n\to\infty} L(g, n, \delta) = L(g, x).$$

The necessary and sufficient condition for the existence of the function $F(x)$ then becomes

$$U(g, x) + L(g, x) = K(x),$$

where K is independent of g. The functions $U(g, x)$ and $L(g, x)$ are monotone in x, and bounded. Consequently $K(x)$ is of bounded variation, and this shows that

$$F(x) = \int_a^x f\,dx + K(x)$$

is of bounded variation, a fact which can easily be proved independently of the foregoing. We now prove

THEOREM IV. *If $F(x)$ is a function of bounded variation on (a, b), then there exists a function $f(x)$ and a sequence of summable functions $s_n(x)$ tending to $f(x)$ with $\int_e s_n(x)dx$ bounded in n and e such that $\int_a^x s_n(x)dx$ tends to $F(x)-F(a)$.*

If d_i is a discontinuity of F, set

$$s_l(d_i) = F(d_i) - F(d_i - 0), \qquad s_r(d_i) = F(d_i + 0) - F(d_i),$$

$$\phi(x) = \sum_{a\leqq d_i<x} s_r(d_i) + \sum_{a\leqq d_i\leqq x} s_l(d_i).$$

Then $F(x)=\phi(x)+\psi(x)$ where $\psi(x)$ is continuous.* Hence given $\epsilon>0$ there exists $\delta>0$ such that if x' and x'' are any two points on (a, b) with $|x'-x''| <\delta$ then

$$(1) \qquad |F(x') - F(x'')| < \sum_{(a,b)} |s_l(d_i)| + \sum_{(a,b)} |s_r(d_i)| + \epsilon.$$

Arrange the discontinuities of F in a definite order $d_1, d_2, \cdots$, and consider the intervals A_n defined on (a, b) by the points $d_1, \cdots, d_n$, where

$$(2) \qquad \sum_{i=n+1}^{\infty} |s_l(d_i)| + \sum_{i=n+1}^{\infty} |s_r(d_i)| < \epsilon.$$

Let (a, d) be the first of these intervals and let $F_1' = F'$, where F' is finite, $F_1' = 0$ elsewhere. On this interval (a, d) there is a set of points E with $mE = d-a$ at which (i) F is continuous; (ii) F' exists; (iii) $\int_x^{x+h} F_1'\,dx/h$ tends to F' as h tends to zero; (iv) each point of E is a point of density of E. From (iii) we get, for h sufficiently small,

* Lebesgue, *Leçons sur l'Intégration*, Paris, 1928, p. 61.

$$\left| \frac{F(x+h) - F(x)}{h} - \frac{1}{h}\int_x^{x+h} F_1' \, dx \right| < \epsilon. \tag{3}$$

From this it follows that for each x of E there exists $\delta' < \delta$, δ fixed in (1) above, for which

$$\left| F(\xi) - F(x) - \int_x^\xi F_1' \, dx \right| < \epsilon(\xi - x), \qquad (x \leqq \xi \leqq x + \delta'), \tag{4}$$

where, on account of (iv), $x+\delta'$ can be taken as a point of E. Consequently with each x of E there is associated an infinite sequence of intervals $(x, x+\delta_i')$ for which (3) holds with δ_i' tending to zero, and with $x+\delta_i'$ points of E. It is, therefore, possible to select a finite non-overlapping set (x_i, x_{i+1}) of these intervals for which (4) holds, and for which

$$\sum (x_{i+1} - x_i) > d - a - \eta, \qquad \eta < \delta. \tag{5}$$

Let (x_j, x_{j+1}) be the intervals on (a, d) complementary to the set (x_i, x_{i+1}). Order the intervals of these two sets from left to right into the set (x_k, x_{k+1}). On the intervals $(x_0 = a, x_1)$, $(x_{l-1}, x_l = d)$, and the remaining intervals of (x_k, x_{k+1}) which belong to the set (x_j, x_{j+1}) let $s_\epsilon(x) = \{F(x_{k+1}) - F(x_k)\}/(x_{k+1} - x_k)$. On the remaining intervals of the set (x_k, x_{k+1}) let $s_\epsilon(x) = F'$, where F' is finite. Otherwise let $s_\epsilon(x) = 0$. For a fixed $k = k'$ other than $k = 0, l-1$ it follows from (1), (2), (4), and (5) that

$$\left| F(x) - F(x_{k'}) - \int_{x_{k'}}^x s_\epsilon dx \right| \qquad (x_{k'} \leqq x \leqq x_{k'+1}) \tag{6}$$

is not greater than the greater of the two numbers 2ϵ and $\epsilon(x - x_{k'})$. We have

$$\int_a^x s_\epsilon dx = \int_a^{x_{k'}} s_\epsilon dx + \int_{x_{k'}}^x s_\epsilon dx$$
$$= \int_a^{x_1} s_\epsilon dx + \sum \int_{x_j}^{x_{j+1}} s_\epsilon dx + \sum \int_{x_i}^{x_{i+1}} s_\epsilon dx + \int_{x_{k'}}^x s_\epsilon dx,$$

where in each case the sum is taken over the intervals of the sets (x_j, x_{j+1}), (x_i, x_{i+1}) to the left of $x_{k'}$ except (x_0, x_1). It now follows from (4), (6), and the definition of s_ϵ, that

$$\left| F(x) - F(a) - \int_a^x s_\epsilon dx \right| < \epsilon(x - a) + 2\epsilon, \tag{7}$$

for $x_1 \leqq x \leqq x_{l-1}$, and for $x = d$. Also $s_\epsilon = F'$ on a set e with $me > d - a - 2\eta$. This construction can be repeated for each of the intervals of the set A_n in such a way that relation (7) holds for each point of (a, b) except possibly the points

interior to a set of intervals $\alpha_n=(a, a+\delta_{n0})$, $(d_i-\delta_{ni}, d_i)$, $(d_i, d_i+\delta_{ni})$, $(b-\delta_{n0}', b)$, where d_i represents the points of $d_1, \cdots, d_n$ other than a and b. Furthermore s_ϵ, now defined on (a, b), is such that $s_\epsilon=F'$ on a set E_ϵ with $mE_\epsilon>b-a-2n\eta$, where η is arbitrarily small independently of n. If ϵ_n is a sequence of values of ϵ tending to zero, then for the corresponding sequence of functions $\sigma_n=s_{\epsilon_n}$ it will now be shown that

$$F(x) - F(a) = \lim_{n\to\infty}\int_a^x \sigma_n dx, \tag{8}$$

for all values of x on (a, b). Relation (7) holds for the discontinuities $d_1, \cdots, d_n$. Consequently (8) holds for all the discontinuities of F. There remains the consideration of points of continuity of F which are on an infinite set of the open intervals α_n. For a fixed d_i the intervals $(d_i-\delta_{ni}, d_i)$, $(d_i, d_i+\delta_{ni})$ are such that δ_{ni} tends to zero as n increases. Hence if any point x is on the first of these open intervals for an infinite set of values of n, then

$$\left| F(x) - F(d_i - \delta_{ni}) - \int_{d_i-\delta_{ni}}^x \sigma_n dx \right| < \sum_{i=n+1}^{\infty} |s_l(d_i)| + \sum_{i=n+1}^{\infty} | s_r(d_i) | + | \psi(x) - \psi(d_i - \delta_{ni}) |.$$

As n increases each of the three terms on the right tends to zero. A similar relation holds for x on $(d_i, d_i+\delta_{ni})$. This, with the foregoing, establishes (8) for every point x of (a, b). The sequence $\sigma_n=F'$ on E_n where mE_n tends to $b-a$. It then follows that there exists a subsequence s_n of σ_n and a set $\mathcal{E}$ with $m\mathcal{E}=b-a$ such that s_n tends to F' on $\mathcal{E}$. Let $f(x)=F'$ on $\mathcal{E}$, $f(x)=0$ elsewhere on (a, b). Then this sequence s_n tends to f almost everywhere, and (8) holds with s_n replacing σ_n. From the manner in which s_n was constructed, it is clear that $\int_e s_n dx$ is bounded in n and e. The function f and the sequence s_n satisfy the requirements of the theorem.

That $\int_e s_n dx$ be bounded in n and e is a sufficient condition for F to be of bounded variation, but it is not a necessary condition. Let $x_1=a<x_2<x_3<\cdots$ be a sequence of values of x on (a, b) with x_n tending to b. On the intervals (x_{n-1}, x_n), (x_n, x_{n+1}) let s_n be constant and such that the integrals of s_n over these intervals is n and $-n$ respectively. Then $F(x)=0$, $a\leqq x\leqq b$, but $\int_e s_n$ is not bounded in n and e.

3. **The independence of $F(x)$ and $\int_a^x f(x)dx$.** Some examples are now given which show that the limit function $F(x)$ is independent of $f(x)$. We first construct a special sequence $s(a, b, r)$ on the linear interval (a, b). Delete the interior points of the middle third of (a, b), then the interior points of the middle third of each of the remaining thirds, and so on indefinitely. Let G be

the non-dense closed set of zero measure which remains. At the nth stage of this process there are 2^n undeleted intervals. On each of these intervals let $s_n(x)$ be constant, and such that $\int s_n dx$ over each interval is equal to $r/2^n$, where r is a prescribed real number not zero. Let $s_n=0$ elsewhere on (a, b). If s_n is now redefined to be zero at the points of G, then $\int s_n$ over any part of (a, b) is not changed, and s_n tends to zero everywhere. Furthermore, it is easily verified that

$$\phi(x) = \lim_{n\to\infty} \int_a^x s_n dx$$

is continuous, monotone, and $\phi(b)-\phi(a)=r$. Let this sequence s_n be denoted by $s(a, b, r)$.

Let $F(x)$ be any continuous function on (a, b). Divide (a, b) into n equal parts by the points $a=d_0, d_1, \cdots, d_n=b$. Let $F(d_i)-F(d_{i-1})=r_i$. On (d_{i-1}, d_i) let s_n be the kth member of the sequence $s(d_{i-1}, d_i, r_i)$, where k is sufficiently great to insure that $s_n=0$ on a part of (a, b) with measure greater than $b-a-1/2^n$. Then $\int_a^{d_i} s_n dx=F(d_i)-F(a)$. Since on (d_i, d_{i+1}), $\int_{d_i}^x s_n dx$ is monotone, for x different from d_i, $\int_a^x s_n dx$ does not differ from $F(x)-F(a)$ by more than the maximum of $|F(x)-F(d_i)|$, $d_i \leqq x \leqq d_{i+1}$. Since this maximum tends to zero as n increases, it follows that

$$F(x) - F(a) = \lim_{n\to\infty} \int_a^x s_n dx. \tag{1}$$

Since $s_n=0$ on a set E_n with $mE_n>b-a-1/2^n$, it follows that there exists a set E on (a, b) with $mE=b-a$ at each point of which s_n tends to zero. If s_n is now redefined to be zero at the points of CE, then (1) holds with s_n tending to zero everywhere. Now let $x_1, x_2, \cdots$ be a sequence of values of x with $a< \cdots <x_n<x_{n-1}< \cdots$ and x_n tending to a. Redefine s_n on the interior of (a, x_n) in such a way that $\int_a^{x_n} s_n dx=F(a)$. We then have

$$F(x) = \lim_{n\to\infty} \int_a^x s_n dx \tag{2}$$

with s_n tending to zero everywhere. We have thus shown:

If $F(x)$ is any continuous function on (a, b), then there exists $S(0, a, x)$ such that $F(x)=S(0, a, x)$, and the sequence s_n used in defining $S(0, a, x)$ tends to zero everywhere.

Next let $F(x)$ be any measurable function on (a, b). There exists a sequence of continuous functions $\phi_n(x)$ tending to $F(x)$ almost everywhere. By the foregoing there exists a sequence s_{nk} such that $s_{nk}=0$ on a set E_{nk} with $mE_{nk}>b-a-\epsilon_k$ and

$$\left| \phi_n(x) - \int_a^x s_{nk} dx \right| < \epsilon_k.$$

The quantity ϵ_k is independent of n. Hence if $\epsilon_k < 1/2^n$ and $s_n = s_{nk}$, then almost everywhere on (a, b)

$$F(x) = \lim_{n\to\infty} \int_a^x s_n dx, \tag{3}$$

and s_n tends to zero almost everywhere. By modifying s_n on a set of zero measure we have (3) holding with s_n tending to zero everywhere. We thus get:

If $F(x)$ is any measurable function on (a, b), then there exists $s_n(x)$ tending to zero everywhere for which $\int_a^x s_n dx$ tends to $F(x)$ almost everywhere.

Finally, let $F_1(x), F_2(x), \cdots$ be any sequence of measurable functions on (a, b). There exists s_{nk} tending to zero everywhere for which

$$F_n(x) = \lim_{k\to\infty} \int_a^x s_{nk} dx$$

almost everywhere. If from the double sequence s_{nk} there is selected the single sequence $s_{1k_1}, s_{1k_2}, s_{2k_2}, s_{1k_3}, s_{2k_3}, s_{3k_3}, \cdots$ then for the single sequence s_n obtained in this way the set of limits of $\int_a^x s_n dx$ includes the sequence of functions $F_1(x), F_2(x), \cdots$ for almost all points of (a, b). If in defining s_n each successive k_p is chosen sufficiently great then s_n tends to zero almost everywhere. By redefining this sequence s_n at a set of zero measure we have the following:

If $F_1(x), F_2(x), \cdots$ is any sequence of measurable functions on (a, b), there exists a sequence of summable functions $s_n(x)$ tending to zero everywhere such that for almost all points of (a, b) the set of limits of $\int_a^x s_n dx$ includes the sequence $F_1(x), F_2(x), \cdots$.

4. **The inversion of approximate derived numbers.** If $f(x)$ is finite except for a denumerable set, and almost everywhere is equal to one* of the derived numbers of the continuous function $F(x)$, then f is integrable in the Denjoy sense to $F(x) - F(a)$. We now obtain the corresponding theorem for approximate derived numbers, with a set of measure zero replacing the denumerable set, and sequence integration replacing Denjoy integration.

Let $f(x)$ be measurable and finite almost everywhere on (a, b), and almost everywhere be equal to one of the approximate derived numbers of the continuous function $F(x)$. Since F is continuous it is measurable. Then, since f is finite except for a set of zero measure, it follows that $f = ADF$ almost every-

* Not necessarily the same derived number at each point.

where.* Let E_n be the set for which $-n<f<n$. Then f is summable over E_n, at almost all points of E_n the density of E_n is unity, and

$$\lim_{h\to 0}\frac{1}{h}\int_{E_n(x,\,x+h)} f\,dx = f. \tag{1}$$

At a point x for which (1) holds let E_x be any measurable set with right-hand density unity at x. Then

$$\lim_{h\to 0}\frac{1}{h}\int_{E_xE_n(x,\,x+h)} f\,dx = f.$$

For if $E'=E_n(x,\,x+h)-E_xE_n(x,\,x+h)$, then

$$\frac{1}{h}\int_{E_n(x,\,x+h)} f\,dx = \frac{1}{h}\int_{E_xE_n(x,\,x+h)} f\,dx + \frac{1}{h}\int_{E'} f\,dx.$$

Since $|f|<n$ on E_n, the second integral on the right is at most equal to nmE'/h, and as h tends to zero this tends to zero for the reason that the density of E' is zero at x. Since at almost all points of E_n, f is the approximate derivative of F, we have for these points x,

$$\lim_{\xi\to x}\frac{F(\xi)-F(x)}{\xi-x} = f, \qquad \xi > x, \tag{2}$$

for a set e_ξ of right-hand density unity at x. Hence, from (1) and (2), for a given $\epsilon>0$, and a given η with $0<\eta<1$, there exists $\delta>0$ for which

$$\left|\frac{F(\xi)-F(x)}{\xi-x} - f\right| < \frac{\epsilon}{2}, \qquad \left|\frac{1}{h}\int_{e_\xi(x,\,x+h)} f\,dx - f\right| < \frac{\epsilon}{2},$$

$me_\xi(x,\,x+h)>\eta h$, $0<h<\delta$, δ depending on x, and this relation holding for almost all points of E_n. For n sufficiently large mE_n is arbitrarily near to $b-a$. We conclude, therefore, that there exists on $(a,\,b)$ a set E with $mE=b-a$, to each point x of which there corresponds a set e_x and a number $\delta>0$ for which,

$$\left|\frac{F(\xi)-F(x)}{\xi-x} - \frac{1}{\xi-x}\int_{e_x(x,\xi)} f\,dx\right| < \epsilon,$$

ξ belonging to e_x, $me_x(x,\,x+h)>\eta h$, $0<h<\delta$. Hence to each point x of E there corresponds a sequence of intervals $(x,\,x+h_i)$ with h_i tending to zero

* J. C. Burkill, and U. S. Haslam-Jones, Proceedings of the London Mathematical Society, ser. 2, vol. 32 (1900), pp. 346–355. It is shown that if F is measurable and finite, then almost everywhere ADF exists and is finite, or $AD^+=AD^-=\infty$, $AD_+=AD_-=-\infty$.

such that on $(x, x+h_i)$ there is a set e_x which includes the point $x+h_i$, with $me_x(x, x+h_i)>\eta h_i$, and for which

$$\left| F(\xi) - F(x) - \int_{e_x(x,\xi)} f\,dx \right| < \epsilon(\xi - x), \tag{4}$$

where on $(x, x+h_i)$ the set e_ξ is the set $e_x(x, x+h_i)$. From the set of intervals thus associated with the set E, it is possible to select a finite set (x_k, x_{k+1}) with $\sum(x_{k+1}-x_k)>mE-\epsilon$. Let (x_j, x_{j+1}) be the intervals on (a, b) complementary to the set (x_k, x_{k+1}). Furthermore, let (x_k, x_{k+1}) be so chosen that if x' and x'' are any two points on an interval of the set (x_k, x_{k+1}) or on an interval of the set (x_j, x_{j-1}) then

$$\left| F(x') - F(x'') \right| < \epsilon. \tag{5}$$

Let $s_{\epsilon\eta}=f$ on $e_{x_k}=e_k$, $s_{\epsilon\eta}=\{F(x_{j+1})-F(x_j)\}/(x_{j+1}-x_j)$ on (x_j, x_{j+1}), and $s_{\epsilon\eta}=0$ elsewhere on (a, b). The function $s_{\epsilon\eta}=f$ on $\sum e_k$, which is a set with measure $>\eta(b-a-\epsilon)$. For $k=k'$ and x a point of $e_{k'}$ we have

$$\begin{aligned}\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}dx \right| \leqq{}& \left| F(x) - F(x_{k'}) - \int_{x_{k'}}^x s_{\epsilon\eta}dx \right| \\ &+ \sum \left| F(x_{i+1}) - F(x_i) - \int_{x_i}^{x_{i+1}} s_{\epsilon\eta}dx \right| \\ &+ \sum \left| F(x_{l+1}) - F(x_l) - \int_{x_l}^{x_{l+1}} s_{\epsilon\eta}dx \right|,\end{aligned}$$

where (x_i, x_{i+1}) includes all the intervals of (x_k, x_{k+1}) with $k<k'$ and (x_l, x_{l+1}) includes all the intervals of the set (x_j, x_{j+1}) to the left of $x_{k'}$. On account of (4) the first term on the right is not greater than $\epsilon(x-x_{k'})$, and the second term is not greater than $\epsilon(x_{k'}-a)$. While from the definition of $s_{\epsilon\eta}$ on (x_j, x_{j+1}), the third term is zero. Hence for x any point of the set $\sum e_k$,

$$\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}dx \right| < \epsilon(x - a). \tag{6}$$

Let e_k' be a closed subset of e_k for which

$$\left| \int_{e_k - e_k'} f\,dx \right| < \epsilon, \tag{7}$$

and let (α, β) be an interval on (x_k, x_{k+1}) complementary to e_k'. If x is any point on (α, β) which is not a point of e_k then it follows from (5), (6), and (7) that

$$\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}dx \right| < \epsilon(x - a) + 2\epsilon.$$

On an interval of the set (x_i, x_{i+1}), $s_{\epsilon\eta}$ is constant. Consequently the integral of $s_{\epsilon\eta}$ over (x_i, x) is linear on (x_i, x_{i+1}) and varies from zero to $F(x_{i+1})-F(x_i)$. It then follows from the relation (5) that if x is a point on the interval (x_i, x_{i+1}), we have

$$\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}dx \right| < \epsilon(x - a) + 2\epsilon.$$

Hence for all values of x on (a, b),

$$\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}dx \right| < \epsilon(x - a) + 2\epsilon,$$

and $s_{\epsilon\eta}=f$ on a set with measure $>\eta(b-a-\epsilon)$. If then we take a sequence of values of ϵ tending to zero and a corresponding sequence of values of η tending to unity, we arrive at a sequence of functions, s_n, for which

$$F(x) - F(a) = \lim_{n\to\infty} \int_a^x s_n dx.$$

Furthermore, since $s_n=f$ on a set with measure $>\eta_n(b-a-\epsilon_n)$ it follows that there exists a subsequence of s_n which converges to f almost everywhere on (a, b). We have thus proved

THEOREM V. *Let the function $f(x)$ be measurable and finite almost everywhere on (a, b), and almost everywhere be one or the other of the approximate derived numbers of the continuous function $F(x)$. Then $f(x)$ is integrable in the sequence sense to $F(x)-F(a)$.*

5. **The limit of $s_n(x)$ not summable.** Let $x_1, x_2, \cdots$ be a sequence of values of x on $(0, 1)$ with $x_1=0$, $x_i<x_{i+1}$, and with x_i tending to unity. On (x_i, x_{i+1}) let $f=\pm 1/[i(x_{i+1}-x_i)]$, $+$ or $-$ holding accordingly as i is odd or even. Let $s_n=f$ on (x_i, x_{i+1}), $i=1, 2, \cdots, n$, and $s_n=0$ elsewhere. Then the integral of f exists on $(0, 1)$ as a non-absolutely convergent integral, s_n is summable for each n, and

$$\int_0^x f(x)dx = F(x) = \lim_{n\to\infty} \int_0^x s_n(x)dx, \qquad 0 \leqq x \leqq 1. \tag{1}$$

Let $\delta_i=(x_i, x_{i+1})$. On δ_i the function f is positive or negative accordingly as i is odd or even. Let the sequence $\delta_1, \delta_2, \cdots$ be rearranged in the order $\delta_1, \delta_3, \delta_2, \delta_5, \delta_7, \delta_4, \cdots$, two intervals on which f is positive followed by

one on which f is negative. Let this rearranged sequence be $\gamma_1, \gamma_2, \cdots$, let $s_n=f$ on $E_n=\gamma_1, \gamma_2, \cdots, \gamma_n$, and let $s_n=0$ elsewhere on $(0, 1)$. Then

$$(2)\quad \lim_{n\to\infty}\int_0^x s_n dx = \int_0^x f(x)dx, \quad 0 \leqq x < 1; \quad \lim_{n\to\infty}\int_0^1 s_n dx > \int_0^1 f\,dx.$$

In both (1) and (2) $s_n=f$ on E_n, $E_n \supset E_{n-1}$, and mE_n tends to unity. We thus see that for the function f defined above there is a sequence s_n of this general type for which (1) holds, and another such sequence for which (1) does not hold. This raises the question: If f is any function which is integrable in a non-absolutely convergent sense, does there exist at least one sequence of this general type for which (1) holds? The answer is in the affirmative for the generalized Denjoy integral. The proof of this is built up in several stages.

Let $f(x)$ be a measurable function which is integrable in the generalized Denjoy sense on (a, b). Let E_1 be the points of non-summability of f on (a, b), and (α_i, β_i) the set of open intervals complementary to E_1. Fix ϵ_n and let (α_n', β_n') be an interval with $\alpha_i<\alpha_n'<\beta_n'<\beta_i$, and such that for x on (α_i, α_n'), (β_n', β_i) we have respectively

$$|F(x) - F(\alpha_i)| < \epsilon_n, \qquad |F(x) - F(\beta_i)| < \epsilon_n.$$

The function f is summable on (α_n', β_n'), and if $s_n=f$ on (α_n', β_n') and $s_n=0$ elsewhere on (α_i, β_i), then

$$\left| F(x) - F(\alpha_i) - \int_{\alpha_i}^x s_n dx \right| < 2\epsilon_n, \qquad \alpha_i \leqq x \leqq \beta_i.$$

Letting ϵ_n tend to zero and α_n', β_n' tend monotonically to α_i, β_i respectively we get

$$F(x) - F(\alpha_i) = \lim_{n\to\infty}\int_{\alpha_i}^x s_n dx, \qquad \alpha_i \leqq x \leqq \beta_i,$$

where $s_n=f$ on E_n, $s_n=0$ elsewhere on (α_i, β_i), $E_n \supset E_{n-1}$, and mE_n tends to $\beta_i-\alpha_i$. Hence we have:

On each interval (α_i, β_i) *complementary to* E_1 *there exists* $TS(f, a, x)$ *such that* $F(x)-F(\alpha_i)=TS(f, a, x)$.

Let E_2 be the points of non-summability of f over E_1 together with the points of E_1 at which $\sum|F(\beta_i)-F(\alpha_i)|$ diverges, (α_i, β_i) the intervals complementary to E_1. Now let (α_i, β_i) be an interval of the set complementary to E_2. Fix ϵ_n and let (α_n', β_n') be an interval with $\alpha_i<\alpha_n'<\beta_n'<\beta_i$ such that for x on (α_i, α_n'), (β_n', β_i) we have respectively

$$(1) \qquad |F(x) - F(\alpha_i)| < \epsilon_n, \qquad |F(x) - F(\beta_i)| < \epsilon_n.$$

Let e be the part of E_1 on (α_n', β_n'), (α_j, β_j) the intervals complementary to e on (α_n', β_n'). Then $\sum |F(\beta_j) - F(\alpha_j)|$ converges. Fix p_n so that

$$(2) \qquad \sum_{j=p_n+1}^{\infty} |F(\beta_j) - F(\alpha_j)| < \epsilon_n,$$

and so that for $j > p_n$ and x on (α_j, β_j),

$$(3) \qquad |F(x) - F(\alpha_j)| < \epsilon_n.$$

At a point x of e on (α_n', β_n') we have

$$(4) \qquad F(x) - F(\alpha_n') = \sum_{(\alpha_n', x)} \{F(\beta_j) - F(\alpha_j)\} + \int_{e(\alpha_n', x)} f\,dx.$$

On (α_j, β_j) $(j = 1, 2, \cdots, p_n)$ there exists E_{nj} with mE_{nj} arbitrarily near to (α_j, β_j), and $s_{nj} = f$ on E_{nj}, $s_{nj} = 0$ elsewhere on (α_j, β_j) for which

$$(5) \qquad \left| F(x) - F(\alpha_j) - \int_{\alpha_j}^{x} s_{nj}\,dx \right| < \frac{\epsilon_n}{p_n}, \qquad \alpha_j \leqq x \leqq \beta_j.$$

If $s_n = s_{nj}$ on (α_j, β_j), $s_n = f$ on e, and $s_n = 0$ elsewhere on (α_i, β_i), it follows from (1), (2), (3), (4), and (5) that

$$\left| F(x) - F(\alpha_i) - \int_{\alpha_i}^{x} s_n\,dx \right| < 4\epsilon_n.$$

Let ϵ_n tend to zero, α_n', β_n' tend monotonically to α_i, β_i respectively, and p_n increase monotonically. Also on the intervals (α_j, β_j), $j = 1, 2, \cdots, p_n$, let E_{nj} be so determined that E_{nj} contains $E_{(n-1)j}$ and $\sum_j mE_{nj}$ tends to $\sum(\beta_j - \alpha_j)$. Then

$$F(x) - F(\alpha_i) = \lim_{n\to\infty} \int_{\alpha_i}^{x} s_n\,dx, \qquad \alpha_i \leqq x \leqq \beta_i,$$

$s_n = f$ on E_n, $E_n \supset E_{n-1}$ and mE_n tends to $\beta_i - \alpha_i$. Thus we have:

On all the intervals (α_i, β_i) complementary to E_2 there exists $TS(f, a, x)$ for which $F(x) - F(\alpha_i) = TS(f, a, x)$, $\alpha_i \leqq x \leqq \beta_i$.

If E_3 is the set of points of non-summability of f over E_2 together with the points of E_2 at which $\sum |F(\beta_i) - F(\alpha_i)|$ diverges, (α_i, β_i) the intervals complementary to E_2, then the foregoing process can be repeated to obtain the corresponding result for the intervals (α_i, β_i) complementary to E_3. Furthermore, the process can be repeated for every set E_λ for $\lambda < \omega$, where ω is the

first transfinite ordinal of the second kind. Let (α_i, β_i) be an interval of the set complementary to E_ω and (α_n', β_n') an interval with $\alpha_i<\alpha_n'<\beta_n'<\beta_i$ and with

$$(6) \qquad |F(x) - F(\alpha_i)| < \epsilon_n, \qquad |F(x) - F(\beta_i)| < \epsilon_n,$$

for x on (α_i, α_n'), (β_n', β_i) respectively. There is a set E_λ, $\lambda<\omega$ for which the part of E_λ on (α_n', β_n') is empty. It then follows that the methods of construction given above lead to the existence of $TS(f, \alpha_n', x)$, for which $TS(f, \alpha_n', x) = F(x)-F(\alpha_n')$, $\alpha_n' \leqq x \leqq \beta_n'$. Hence on (α_n', β_n') there exists $s_{nk}=f$ on E_{nk}, $s_{nk}=0$ elsewhere on (α_n', β_n'), $mE_{nk}>\beta_n-\alpha_n-\epsilon_n$, and

$$(7) \qquad \left| F(x) - F(\alpha_n') - \int_{\alpha_n}^{x} s_{nk}dx \right| < \epsilon_n, \qquad \alpha_n' \leqq x \leqq \beta_n'.$$

If ϵ_n tends to zero and α_n', β_n' tend respectively to α_i, β_i, it follows from (6) and (7) that, if $s_{nk}=s_n$, then

$$F(x) - F(\alpha_i) = \lim_{n\to\infty} \int_{\alpha_i}^{x} s_n dx, \qquad \alpha_i \leqq x \leqq \beta_i,$$

where $s_n=f$ on E_n, $s_n=0$ elsewhere on the interval (α_i, β_i) and mE_n tends to $\beta_i-\alpha_i$. Let $\epsilon_1, \epsilon_2, \cdots$ be a sequence of values of ϵ tending to zero. Fix n_1 so that

$$(8) \qquad \left| F(x) - F(\alpha_i) - \int_{\alpha_1}^{x} s_{n_1}dx \right| < \epsilon_1, \qquad \alpha_i \leqq x \leqq \beta_i.$$

Let $\mathcal{E}_1=E_{n_1}$ and $\sigma_1=s_{n_1}$. Fix n_2 so that

$$(9) \qquad \left| F(x) - F(\alpha_i) - \int_{\alpha_i}^{x} s_{n_2}dx \right| < \epsilon_2, \qquad \alpha_i \leqq x \leqq \beta_i,$$

and so that

$$(10) \qquad \int_G |f|\, dx < \epsilon_2,$$

where $G=\mathcal{E}_1-E_{n_2}$. Relation (10) is possible for the reason that f is summable on $\mathcal{E}_1$ and mE_n tends to $(\beta_i-\alpha_i)$. Set $\mathcal{E}_2=E_{n_2}+(\mathcal{E}_1-\mathcal{E}_{n_2})$. Then $\mathcal{E}_2\supset\mathcal{E}_1$. Also, if $\sigma_2=f$ on $\mathcal{E}_2$, $\sigma_2=0$ elsewhere on (α_i, β_i) it follows from (9) and (10) that

$$\left| F(x) - F(\alpha_i) - \int_{\alpha_i}^{x} \sigma_2 dx \right| < 2\epsilon_2, \qquad \alpha_i \leqq x \leqq \beta_i.$$

This process can be repeated indefinitely, giving a sequence of functions $\sigma_1, \sigma_2, \cdots$ for which

$$F(x) - F(\alpha_i) = \lim_{n\to\infty} \int_{\alpha_i}^{x} \sigma_n dx,$$

where $\sigma_n = f$ on $\mathcal{E}_n$, $\sigma_n = 0$ elsewhere on (α_i, β_i), $\mathcal{E}_n \supset \mathcal{E}_{n-1}$ and $m\mathcal{E}_n$ tends to $\beta_i - \alpha_i$. This allows us to state:

If ω is the first transfinite ordinal of the second kind and (α_i, β_i) an interval on (a, b) complementary to E_ω, then there exists $TS(f, \alpha_i, x)$ for which $TS(f, \alpha_i, x) = F(x) - F(\alpha_i)$, $\alpha_i \leqq x \leqq \beta_i$.

The processes of construction given above can be repeated to give the corresponding result for an interval (α_i, β_i) of the set complementary to E_λ, where λ is any finite or transfinite ordinal of the first kind, or for an interval (α_i, β_i) of the set complementary to E_ω, where ω is any transfinite ordinal of the second kind. The method of transfinite induction can now be used to prove:

THEOREM VI. *If $f(x)$ is measurable on (a, b) and integrable in the generalized Denjoy sense to $F(x)$, then there exists $TS(f, a, x)$ for which $TS(f, a, x) = F(x)$, $a \leqq x \leqq b$.*

The sequence $s_1, s_2, \cdots$ of Theorem V converges to f almost everywhere, but is not defined wholly in terms of f. There may be intervals (x_j, x_{j+1}) on (a, b) for which $s_n = \{F(x_{j+1}) - F(x_j)\}/(x_{j+1} - x_j)$, and consequently s_n cannot be determined on these intervals without a knowledge of the values of F at the points x_j, x_{j+1}. In some cases s_n can be determined without a knowledge of the values of F at particular points. To throw further light on this point we start with a continuous function $F(x)$ which is also (ACG), and prove

LEMMA I. *Let $F(x)$ be (ACG) on (a, b), and let e be any closed set with $me = 0$. Then there exists a finite set of intervals (a_n, b_n) with (a_n, b_n) points of e which contain, either as end points or interior points, all of e except at most a finite set, and for which*

$$\sum |F(b_n) - F(a_n)| < \epsilon, \qquad \sum (b_n - a_n) < \delta,$$

where ϵ and δ are arbitrary positive numbers.

Under the conditions of the lemma $e = e_1 + e_2 + \cdots$, where F is absolutely continuous on each e_n. Let $\bar{e}_n$ be the set e_n together with its limit points. Then, since e is closed, $\bar{e}_n \subset e$, and the continuity of F can be used to show that F is absolutely continuous on $\bar{e}_n$. Let (α_i, β_i) be the intervals on (a, b) contiguous to $\bar{e}_1$, and (c_{1i}', c_{1i}'') the finite set of intervals on (a, b) belonging to the complement of the intervals $(\alpha_1, \beta_1), \cdots, (\alpha_n, \beta_n)$. Then the intervals (c_{1i}', c_{1i}'') contain, either as end points or interior points, all of $\bar{e}_1$, except at most a finite set. The points c_i', c_i'' belong to $\bar{e}_1$. Then, since F is

absolutely continuous on $\bar{e}_1$ and since $m\bar{e}_1=0$, it follows that if n is sufficiently great,

$$(1)\qquad \sum \left| F(c_{1i}'') - F(c_{1i}') \right| < \epsilon_1, \qquad \sum (c_{1i}'' - c_{1i}') < \delta_1.$$

Let B_1', B_2', $\cdots$ be the finite set of intervals complementary to the set (c_{1i}', c_{1i}''), and let $\bar{e}_{2i}$ be the part of $\bar{e}_2$ on the closed interval B_i'. There then exists on B_i' a finite set of intervals containing, either as end points or interior points, all of the set $\bar{e}_{2i}$, except at most a finite number of points of $\bar{e}_{2i}$, and which satisfy relations similar to (1). If (c_{2i}', c_{2i}'') is the total set of intervals thus determined for all of the set B_i', it is possible to have

$$(2)\qquad \sum \left| F(c_{2i}'') - F(c_{2i}') \right| < \epsilon_2, \qquad \sum (c_{2i}'' - c_{2i}') < \delta_2.$$

The set $(c_{1i}', c_{1i}'')+(c_{2i}', c_{2i}'')$ contains all of $\bar{e}_1$ and $\bar{e}_2$, except for at most a finite number of points. We designate by B_i^2 the finite set of intervals on (a, b) complementary to $(c_{1i}', c_{1i}'')+(c_{2i}', c_{2i}'')$, and proceed to determine on B_i^2 a finite set of intervals containing all of $\bar{e}_{3i}$, except at most a finite number of points. Continuing this process we arrive at the countable set of intervals, (c_{ji}', c_{ji}'') none of which overlap, which contain, either as end points or as interior points, all of e except at most a countable set P, and for which, if $\sum\epsilon_i<\epsilon$, $\sum\delta_i<\delta$,

$$(3)\qquad \sum_j \sum_i \left| F(c_{ji}'') - F(c_{ji}') \right| < \epsilon, \qquad \sum_j \sum_i (c_{ji}'' - c_{ji}') < \delta.$$

Let D be the end points of the intervals $C=(c_{ji}', c_{ji}'')$, and E the limit points of the set $D+P$. The set $D+P$ is enumerable. Since e is closed and $C+P\supset e$, it follows that $D+P\supset E$. Consequently E is enumerable and closed. Let $E=x_1, x_2, \cdots$. Since e is closed and $me=0$, the set e is non-dense on (a, b). Hence each point x_i of the set E is the left-hand end point of an interval (x_i, x_i'), or the right-hand end point of an interval (x_i', x_i), or both, where in the first case x_i' is a point of e which is not a limit point of e on the right, and in the second case x_i' is a point of e which is not a limit point of e on the left. Thus E is on a set of intervals whose end points are not limit points of e from without. Furthermore, these intervals can be chosen so that

$$(4)\qquad \left| F(x_i') - F(x_i) \right| < \epsilon_i, \qquad \left| x_i' - x_i \right| < \delta_i.$$

Since E is closed, it easily follows that there exists a finite non-overlapping set A of these intervals associated with the set $E=x_1, x_2, \cdots$ which contain, either as end points or interior points, all the points of E. If an end point of the set A happens to be an interior point of an interval (c_{ij}', c_{ij}''), it can be changed to one or the other of the end points of (c_{ij}', c_{ij}'') without altering relations (4). There will then be only a finite set C' of the intervals of C

which are exterior to or abutting the altered set A, and if (a_n, b_n) is the finite set of intervals $C'+A$, then (a_n, b_n) contains, either as interior points or end points, all of e except at most a finite set, the points a_n, b_n belong to e, and

$$\sum |F(b_n) - F(a_n)| < 3\epsilon, \qquad \sum (b_n - a_n) < 3\delta.$$

This establishes the lemma.

Let the function f be finite almost everywhere on (a, b), and almost everywhere be equal to one or the other of the approximate derived numbers of the function F which is (ACG) on (a, b). Working as in Theorem V, it is possible to determine on (a, b) a finite set of intervals (x_k, x_{k+1}), and on each interval a set e_k, where

$$\left| F(\xi) - F(x_k) - \int_{e_k(x_k,\xi)} f\,dx \right| < \epsilon(\xi - x_k), \tag{5}$$

where the set ξ is the set e_k, where $me_k > \eta(x_{k+1}-x_k)$, and where $\sum(x_{k+1}-x_k)$ is arbitrarily near to $b-a$, in particular $>(b-a)/2$. Denote this finite set of intervals by K_1 and the part of (a, b) complementary to K_1 by CK_1. It is then possible to determine on CK_1 a finite set of intervals K_2 satisfying relations similar to (5), with $mK_2 > mCK_1/2$. Denoting the part of (a, b) complementary to K_1+K_2 by CK_2, we can determine a set K_3 on CK_2 with $mK_3 > mCK_2/2$, and with the set K_3 satisfying relations similar to (5). Continuing this process we arrive at a set of non-overlapping intervals $K=K_1+K_2+\cdots$ with $mK=b-a$. Let e be the end points and limit points of end points of K. Then e is closed, $me=0$, and consequently, this set e satisfies the conditions of Lemma I in relation to F. Let (a_n', b_n') be the set of intervals provided by this lemma. It is evident that these intervals can be reduced to a non-abutting set (a_n, b_n) with the same properties relatively to F. The complement of the closed intervals (a_n, b_n) and the finite number of points of e exterior to (a_n, b_n) is a finite set of open intervals (x_k, x_{k+1}) of the set K. Let $s_{\epsilon\eta}=f$ on e_k, $s_{\epsilon\eta}=0$ elsewhere on (a, b). Then, as in Theorem V, for any point of (a, b),

$$\begin{aligned}\left| F(x) - F(a) - \int_a^x s_{\epsilon\eta}\,dx \right| &< \epsilon(x-a) + 2\epsilon + \sum |F(b_n) - F(a_n)| \\ &< \epsilon(b-a) + 3\epsilon.\end{aligned}$$

Also $\sum me_k > \eta\sum(x_{k+1}-x_k) > \eta(b-a-\epsilon)$. If now we take a sequence of values of ϵ tending to zero, and a corresponding sequence of values of η tending to unity, we arrive at a sequence of summable functions s_n for which $s_n=f$ on a set E_n, $s_n=0$ elsewhere, mE_n tends to $b-a$, and

$$F(x) - F(a) = \lim_{n\to\infty} \int_a^x s_n' \, dx. \tag{6}$$

Proceeding as in the concluding part of the proof of Theorem VI, a subsequence σ_n of s_n can be determined for which (6) holds and for which $\sigma_n = f$ on $\mathcal{E}_n$, $\sigma_n = 0$ elsewhere, $\mathcal{E}_n \supset \mathcal{E}_{n-1}$, and $m\mathcal{E}_n$ tends to $b-a$. Thus we have proved

THEOREM VII. *Let $f(x)$ be finite almost everywhere on (a, b), and almost everywhere be equal to one or the other of the approximate derived numbers of the continuous function $F(x)$ which is also (ACG). Then $f(x)$ is totally integrable in the sequence sense to $F(x)-F(a)$.*

If $f(x)$ satisfies the conditions of Theorem VII it is integrable in the generalized Denjoy sense.* Hence Theorem VII follows from Theorem VI. Conversely Theorem VI follows from Theorem VII. For if $F(x)$ is the generalized Denjoy integral of f then F is (ACG), and almost everywhere $ADF = f$. Between the proofs of these two theorems there are these distinctions: Theorem VII holds for continuous functions $F(x)$ which are not (ACG), but which behave relatively to every closed set of zero measure in the manner described by Lemma I, provided such functions exist. Again the proof of Theorem VII does not involve transfinite induction, and gives, therefore, a method for constructing a generalized Denjoy integral without the use of transfinite numbers. This construction is particularly simple when the points E of non-summability of f are of zero measure.

The set E is closed. It then follows from Lemma I that there is a finite set (a_n, b_n) of non-abutting intervals containing all of E except a finite set P, with $\sum |F(b_n) - F(a_n)| < \epsilon$. Let (α_i, β_i) be the finite number of intervals complementary to the set $(a_n, b_n) + P$. If $(\alpha_i^\epsilon, \beta_i^\epsilon)$ is an interval with $\alpha_i < \alpha_i^\epsilon < \beta_i^\epsilon < \beta_i$, then on this interval f is summable. Furthermore, if, for each i, $\alpha_i^\epsilon, \beta_i^\epsilon$ are sufficiently near to α_i, β_i respectively, then

$$\left| F(b) - F(a) - \sum \int_{\alpha_i^\epsilon}^{\beta_i^\epsilon} f \, dx \right| < 2\epsilon,$$

and $F(b) - F(a)$ is obtained by taking a sequence of values of ϵ tending to zero.

6. **Conditions for uniqueness of total sequence integration.** We first construct a function $TS(f, a, x)$ which is not equal to $\int_a^x f(x)dx$. Let G be the Cantor non-dense closed set on $(0, 1)$ defined as in §3. Let (a, b) be the middle third of $(0, 1)$, $x_0 = a < x_1 < x_2 < \cdots$, a sequence of values of x on (a, b) with

* Saks, loc. cit., p. 197, §2.

x_n tending to b. On (x_{i-1}, x_i), $i \neq 1$, let $f(x) = \pm 1/[i(x_i - x_{i-1})]$, $+$ or $-$ holding accordingly as i is odd or even. On (x_0, x_1) let f be constant and such that $\int_{x_0}^{x_1} f\,dx = 1 - \log 2$. Then $\int_a^x f\,dx$ exists as a Denjoy integral on (a, b), with b the single point of non-summability of f, and this integral has the value zero at $x = b$. Let (a', b') be one of the two intervals deleted from $(0, 1)$ in the second step in the construction of G. Let the point x' on $a' \leqq x' \leqq b'$ correspond to the point x on $a \leqq x \leqq b$ by means of a one-to-one projective transformation which carries a' into a and b' into b. At x' on (a', b') let $f(x') = f(x)/2^2$, where x is the point on (a, b) which corresponds to x' on (a', b'). The function $f(x')$ is integrable in the Denjoy sense on (a', b'), b' is the single point of non-summability of $f(x')$, and

$$\int_a^{x'} f(x')dx' = \frac{1}{2}\cdot\frac{1}{2}\int_a^x f(x)dx,$$

x and x' corresponding points on (a', b') and (a, b). On each (a', b') of the four deleted intervals arising in the third step of the construction of G let $f(x')$ be defined similarly, except that $f(x') = f(x)/2^4$, which gives

$$\int_{a'}^{x'} f(x')dx' = \frac{1}{4}\cdot\frac{1}{2^2}\int_a^x f(x)dx,$$

where x and x' are corresponding points. Continuing this process, and setting $f = 0$ on G, the function f is integrable in the Denjoy sense on $(0, 1)$. If (α_i, β_i) are the intervals complementary to G, $F(x) = \int_0^x f\,dx$ is such that $F(\alpha_i) = F(\beta_i) = 0$, and for x a point interior to (α_i, β_i), $F(x) = \int_{\alpha_i}^x f\,dx$.

Let λ be a fixed positive number $> 1 - \log 2$, and on (a, b), the middle third of $(0, 1)$, let $e_1, e_2, \cdots, e_n$ be the intervals $(x_0, x_1), (x_2, x_3), \cdots, (x_{2n}, x_{2n+1})$. On these intervals f is positive. There exists n such that

$$\sum_{i=1}^{n-1}\int_{e_i} f\,dx < \lambda, \qquad \sum_{i=1}^{n}\int_{e_i} f\,dx \geqq \lambda.$$

Let $E_{11} = e_1 + e_2 + \cdots + e_n$. Let $\delta_{21}, \delta_{22}, \delta_{23}$ be the three deleted intervals arising at the end of the second stage in the construction of G, ordered from left to right. On δ_{22} the middle third of (a, b) let $E_{22} = (x_0, x_{2n}) + e_{n+1} + \cdots + e_{n+p}$, where $e_{n+1} = (x_{2n+2}, x_{2n+3})$, $e_{n+2} = (x_{2n+4}, x_{2n+5})$, $\cdots$ and where

$$\int_a^{x_{2n}} f\,dx + \sum_{i=1}^{p-1}\int_{e_{n+i}} f\,dx < \frac{\lambda}{3}, \qquad \int_a^{x_{2n}} f\,dx + \sum_{i=1}^{p}\int_{e_{n+i}} f\,dx \geqq \frac{\lambda}{3}.$$

On δ_{21}, δ_{23} obtain sets E_{21}, E_{23} by repeating the construction of the set E_{11} on (a, b) with $\lambda/3$ replacing λ. Let $\mathcal{E}_1 = E_{11}$, $\mathcal{E}_2 = \sum E_{2i}$. Then $\mathcal{E}_2$ contains $\mathcal{E}_1$. Let $\delta_{31}, \delta_{32}, \cdots, \delta_{37}$ be the seven deleted intervals arising at the end of the

third step in the construction of G. On δ_{3i} construct the set E_{3i} according to the above scheme with $\lambda/7$ replacing $\lambda/3$, taking care that $\mathcal{E}_3=\sum E_{3i}$ contains $\mathcal{E}_2$. This process of construction can be continued, giving the sets $\mathcal{E}_1, \mathcal{E}_2, \cdots$ with $\mathcal{E}_n \supset \mathcal{E}_{n-1}$ and $m\mathcal{E}_n$ tending to unity. Let $s_n=f$ on $\mathcal{E}_n$ and $s_n=0$ elsewhere on $(0, 1)$. Then s_n is summable on $(0, 1)$, s_n tends to f, and it is easily verified that

$$\lim_{n\to\infty}\int_a^x s_n dx = \phi(x), \tag{1}$$

where $\phi(x)$ is continuous. Let x be any point on $(0, 1)$, and let $\mathcal{R}_n(x)$ be the number of whole deleted intervals to the left of x arising at the nth step in the construction of G. There are 2^n-1 of these intervals, and it is easily verified that as n increases $\mathcal{R}_n(x)/(2^n-1)$ tends to a limit $\mathcal{R}(x)$, where $\mathcal{R}(x)$ is continuous, non-decreasing, constant on the intervals (α_i, β_i) complementary to G, and such that $\mathcal{R}(0)=0$, $\mathcal{R}(1)=1$. Furthermore

$$\phi(x) = \lambda\mathcal{R}(x) + \int_a^x f\,dx. \tag{2}$$

We thus have $\phi(x)=TS(f, a, x)\neq\int_a^x f\,dx$.

The function $\phi(x)$ is not (ACG) on $(0, 1)$. Suppose the contrary to be true. If ϕ is (ACG) on $(0, 1)$ then, since G is closed and $mG=0$, it follows from Lemma I that there exists a finite set of intervals (a_n, b_n) on $(0, 1)$ containing all of G except a finite set P, where a_n, b_n are points of G, $\sum(b_n-a_n)$ is arbitrarily small, the intervals (γ_n, δ_n) complementary to $(a_n, b_n)+P$ are a finite number of the intervals (α_i, β_i) complementary to G, and the sum $\sum|\phi(b_n)-\phi(a_n)|$ is arbitrarily small. Since $\mathcal{R}(x)$ is constant on (α_i, β_i), it follows from (2) that $\phi(\beta_i)-\phi(\alpha_i)=\int_{\alpha_i}^{\beta_i}f\,dx=0$. Hence $\phi(\delta_n)-\phi(\gamma_n)=0$. But

$$\phi(1) - \phi(0) = \sum\{\phi(b_n) - \phi(a_n)\} + \sum\{\phi(\delta_n) - \phi(\lambda_n)\}.$$

Since the first term on the right can be made arbitrarily small by a proper choice of the intervals (a_n, b_n) and since the second term on the right is zero for every choice of (a_n, b_n) in accordance with the requirements of Lemma I, it follows that $\phi(1)-\phi(0)=0$. Again, since $\mathcal{R}(0)=0$, $\mathcal{R}(1)=1$, and $\int_0^1 f\,dx=0$, it follows from (2) that $\phi(1)-\phi(0)=1$. We thus get a contradiction, and are able to conclude that $\phi(x)$ is not (ACG) on (a, b).

It is possible to prove:

THEOREM VIII. *If f is integrable in the generalized Denjoy sense on (a, b), if the set E of points of non-summability of f has zero measure, and if $TS(f, a, x)$ exists which is (ACG), then*

$$TS(f, a, x) = \int_a^x f\,dx = F(x).$$

Since $mE=0$ we can consider that $f=0$ on E. There exists an interval (l, m) containing a part e of E on its interior such that if (α_i, β_i) are the intervals on (l, m) complementary to E, then $\sum\{F(\beta_i)-F(\alpha_i)\}$ converges. If (α', β') is an interval such that $\alpha_i<\alpha'<\beta'<\beta_i$, then f is summable on (α', β') and consequently $TS(f, \alpha', \beta')=F(\beta')-F(\alpha')$. It then follows from the continuity of $TS(f, a, x)$ that $TS(f, \alpha_i, \beta_i)=F(\beta_i)-F(\alpha_i)$. Let $\phi(x)=TS(f, a, x)$ and apply Lemma I as above to get (a_n, b_n) and (γ_n, δ_n) with (γ_n, δ_n) an interval of the set (α_i, β_i) and consequently $\phi(\gamma_n)-\phi(\delta_n)=F(\delta_n)-F(\gamma_n)$, and $\sum|\phi(b_n)-\phi(a_n)|$ arbitrarily small. These intervals can be so chosen that $\sum\{F(\delta_n)-F(\gamma_n)\}$ is at the same time arbitrarily near to $F(m)-F(l)$. It follows from these considerations that $\phi(m)-\phi(l)=F(m)-F(l)$, and for x on (l, m), $\phi(x)-\phi(l)=F(x)-F(l)$. The method of transfinite induction can now be used to show that for x on (a, b), $\phi(x)=TS(f, a, x)=F(x)$.

Added in proof, December 9, 1936. It is now known that there exists $TS(f, a, x)$ which is (ACG) with $TS(f, a, x)\neq\int_a^x f\,dx$. The proof of this, together with necessary and sufficient conditions for $TS(f, a, x)=\int_a^x f\,dx$ will be published in a subsequent paper.

ACADIA UNIVERSITY,
WOLFVILLE, NOVA SCOTIA, AND
THE UNIVERSITY OF WISCONSIN,
MADISON, WIS.

TANGENT LINES AND PLANES IN TOPOLOGICAL SPACES*

BY

CHARLES C. TORRANCE

The object of this paper is to extend to topological spaces some of the concepts and theorems of direct infinitesimal geometry. To attain this object we define certain point sets in a topological space which will play the role of lines, planes, and hyperplanes in Euclidean space. By means of these sets we define the terms tangent line, tangent plane, etc. We arrive at our principal results in Theorems 4 and 4a.

Let S be a point space with a system of neighborhoods satisfying

Axiom 1. *There exists a sequence* $\{G_n\}$, $n=1, 2, \cdots$, *of collections of neighborhoods in S such that* (1) *for each n, G_n covers S*; (2) *for each n, G_{n+1} is a subcollection of G_n*; (3) *if R is any neighborhood in S, and if X and Y are points of R, distinct or not, there exists an integer m such that, if g is any neighborhood of the collection G_m containing X, then $\bar{g}$ is a subset of* $(R-Y)+X$; *and* (4) *if H and K are mutually exclusive closed point sets in S and if H is compact, there exists an integer m such that no neighborhood of the collection G_m contains both a point of H and a point of K.*

This axiom consists of the first four parts of R. L. Moore's Axiom 1′† and is sufficient‡ to establish the theorems to be quoted below. In particular, point sequences and their sequential limits have the various properties§ required in

Definition 1. *If $\{A_n\}$ is a sequence of point sets in the space S, then the sequential limiting set of $\{A_n\}$ is the set of all points P such that there exists a sequence $\{P_n\}$ of points with sequential limit P, where P_n is in A_n for each n; the limiting set of $\{A_n\}$ is the set of all points P such that there exist a subsequence $\{n_i\}$ of the sequence $\{n\}$ and a sequence $\{P_{n_i}\}$ of points with sequential limit P, where P_{n_i} is in A_{n_i} for each n_i.*

---

* Presented to the Society, in part, September 12, 1935, under a slightly different title; received by the editors June 9, 1936. The main portion of this paper was presented to the International Congress of Mathematicians at Oslo, July 17, 1936.

† Moore, *Foundations of Point Set Theory*, Colloquium Publications of the American Mathematical Society, vol. 13, New York, 1932, p. 324. Hereafter, we shall designate this reference by the letter M.

‡ M, p. 82.

§ M, Chapter I, Theorems 2, 3, and 11.

REMARK 1. If, in the preceding definition, all the sets A_n have a common subset M, then the limiting set A and the sequential limiting set A' of $\{A_n\}$ each contain $\overline{M}$; if $A_n = M$ for each n, then $A = A' = \overline{M}$.

DEFINITION 2. *If G and H are two collections of point sets in the space S, then G is called H-continuous provided that* (1) *if g is any element of G and $\{g_n\}$ is any sequence of elements of G such that, γ being the sequential limiting set of $\{g_n\}$, the part of γ contained in g is not contained in any element of H, then, $\{B_n\}$ being an arbitrary sequence of points with B_n in g_n for each n, each subsequence of $\{B_n\}$ contains a subsequence having a sequential limit in g; and* (2) *in the preceding notation, each point C of g is the sequential limit of some sequence $\{C_n\}$ of points, where C_n is in g_n for each n.*

REMARK 2. It is apparent that the preceding definition is not essentially altered if merely the sequence $\{B_n\}$ itself is required to contain a subsequence having a sequential limit in g. Nor is the preceding definition essentially altered if each point of g is required to be merely the sequential limit of a sequence $\{C_{n_i}\}$, where $\{n_i\}$ is some subsequence of the sequence $\{n\}$.

LEMMA 1. *If G is an H-continuous collection in S, and if M is a point set in S which is contained in no element of H, then there exists at most one element g of G which contains M.*

Suppose there exist two distinct elements g and g' of G which contain M. Let P be a point of one of these elements, say g', which is not in the other. Let $\{g_n\}$ be the sequence $g', g', g', \cdots$. By Remark 1, M is a subset of the part of the sequential limiting set of $\{g_n\}$ contained in g. Hence g and $\{g_n\}$ are in the relation described in Definition 2. If $\{B_n\}$ is taken to be the sequence $P, P, \cdots$, then P, contrary to hypothesis, is in g.

REMARK 3. If G is an H-continuous collection in S, and if g and h are any two distinct elements of G, then it follows from Lemma 1 that some element of H contains the common part of g and h.

LEMMA 2. *If G is an H-continuous collection in S, and if no element of G is contained in any element of H, then the elements of G are closed and compact.*

Let g be an element of G and let P be a limit point of g. There exists a sequence $\{P_n\}$ of points of g such that $\lim P_n = P$. Let $\{g_n\}$ be the sequence $g, g, \cdots$. Then g and $\{g_n\}$ are in the relation described in Definition 2 since g is contained in no element of H. If $\{B_n\}$ is taken to be $\{P_n\}$, it follows that P is in g. The proof of compactness is similar.

Notation. If H is a collection of point sets, then H^* denotes the point set sum of all the elements of H.

Definition 3. *If G is a collection of point sets such that $G^*=S$, then a subcollection R of G is called an element neighborhood provided that* (1) *R^* is open, and* (2) *if g is an element of G such that $g^* \subset R^*$, then g is an element of R.*

If G is a collection of point sets such that $G^*=S$, then G itself is an element space which may be topologized in the usual manner by means of the element neighborhoods just now defined. Thus it has meaning to speak of a limit element, of a closed set of elements, of an element continuum, etc.

Notation. If G is a collection of point sets, and if $U, V, \cdots$ comprise a collection H of point sets, then $r_{UV\ldots}$ denotes the collection of all those elements g of G such that g contains at least one point of each element of H, and $R_{UV\ldots}$ denotes† the collection of all those elements g of G such that $g^* \subset \overset{*}{r}_{UV\ldots}$.

If in Definition 2, H is the collection of all those point sets which consist of a single point, then G is called point continuous. Moreover, g and $\{g_n\}$ are in the relation described in Definition 2 when, and only when, there exist two sequences $\{A_n\}$ and $\{A_n'\}$ of points with distinct sequential limits in g, where A_n and A_n' are in g_n for each n.

Definition 4. *The elements of a collection $G^{(1)}$ of continua in the space S are called lines if* (1) *$G^{(1)}$ is point continuous*; (2) *if P and Q are distinct points in S, there exists at least one element of $G^{(1)}$ containing P and Q*; and (3) *if U and V are mutually exclusive neighborhoods in S containing the points P and Q respectively, there exist neighborhoods U' and V' of P and Q within U and V such that $r_{U'V'}=R_{U'V'}$, where $r_{U'V'}$ and $R_{U'V'}$ are defined relative to $G^{(1)}$.*

In what follows, all the point sets considered lie in the space S; moreover, it will be supposed that $G^{(1)}$ is a given collection of lines in S and that the word "line" always connotes "line of $G^{(1)}$." By Lemma 1, there exists a unique line containing any two distinct points P and Q; this line will sometimes be denoted by $\overline{PQ}$.

Remark 4. If P is a point, there exists a neighborhood of P which contains no line. To show this, let Q be a point distinct from P. There exists‡ a pair of mutually exclusive neighborhoods U and V of P and Q. If each neighborhood U' of P within U contains a line, then Condition (3) of Definition 4 cannot be met since $U' \cdot V=0$ and $U' \subset \overset{*}{R}_{U'V'}$, where V' is any neighborhood of Q within V. It follows from this remark that every line contains more than one point.

† We use the notation $R_{UV\ldots}$ because of Lemmas 3 and 3a; in what follows, the collection G, relative to which $R_{UV\ldots}$ is defined, is always explicitly understood.

‡ M, Chapter I, Theorem 23.

EXAMPLE 1. In the Euclidean space E_4 of four dimensions, let λ be the family of planes

$$\begin{cases} ax - by + z = m \\ bx + ay + w = n \end{cases}$$

together with the planes

$$\begin{cases} x = \alpha \\ y = \beta. \end{cases}$$

The elements of λ are lines in the sense of Definition 4 when only a bounded convex portion of E_4 is considered, since two points determine an element of λ.

LEMMA 3. *If U and V are mutually exclusive open point sets, then the set of lines R_{UV} is a line neighborhood.*

The proof of this lemma is effected in two stages. Let A be a point not in V and let C be the "cone" of lines containing A and a point of V. Suppose that C^*-A is not open. Then there exists a point P of C^*-A which is the sequential limit of a sequence $\{P_n\}$ of points in the complement of C^*-A. Let Q be a point of the line $\overline{AP}$ in V. By Condition (2) of Definition 2, Q is the sequential limit of a sequence $\{C_n\}$ of points, where C_n is in the line $\overline{AP_n}$ for each n. By Lemma 1, the lines $\overline{AP_n}$ have no point in common with C^*-A. Hence no point C_n is in V. It follows that C^*-A is open.

Now let P be a point of $\overset{*}{R}_{UV}$. There exists a line l containing P, a point of U, and a point of V. Let A be a point of l distinct from P and in either U or V, say U. Let C be the cone of lines containing A and a point of V. Then P is in the open point set C^*-A. But $C^*-A \subset \overset{*}{R}_{UV}$. Hence $\overset{*}{R}_{UV}$ is open. That R_{UV} is a line neighborhood now follows from the definition of the symbol R_{UV}.

Three point sets are called non-collinear when there exists no line containing a point of each of these sets. Three non-collinear point sets are evidently mutually exclusive.

If in Definition 2, H is the collection $G^{(1)}$, then G is called line continuous. Moreover, g and $\{g_n\}$ are in the relation described in Definition 2 when, and only when, there exist three sequences $\{A_n\}$, $\{A_n'\}$, and $\{A_n''\}$ of points with non-collinear sequential limits in g, where A_n, A_n', and A_n'' are in g_n for each n.

DEFINITION 4a. *The elements of a collection $G^{(2)}$ of point continua in S are called planes if* (1) *$G^{(2)}$ is line continuous*; (2) *if P, Q, and R are non-collinear points, there exists at least one element of $G^{(2)}$ containing P, Q, and R*; (3) *if U, V, and W are non-collinear point neighborhoods containing the points P, Q,*

and R, respectively, there exist neighborhoods U', V', and W' of P, Q, and R within U, V, and W such that $r_{U'V'W'}=R_{U'V'W'}$, where $r_{U'V'W'}$ and $R_{U'V'W'}$ are defined relative to $G^{(2)}$; and (4) *if P and Q are distinct points, and if g is an element of $G^{(2)}$ containing P and Q, then $g^* \supset \overline{PQ}^*$.*

In what follows, it will be supposed that $G^{(2)}$ is a given collection of planes in S and that the word "plane" always connotes "plane of $G^{(2)}$." By Lemma 1, there exists a unique plane containing any three non-collinear points P, Q, and R; this plane will sometimes be denoted by $\overline{PQR}$.

REMARK 4a. If l is a line, there exists a line neighborhood R of l such that R^* contains no plane. To show this, let P_1 and P_2 be distinct points of l and let Q be a point not in l. There exists a pair of mutually exclusive open point sets D and D' containing l and Q. It will be shown in Lemma 4 below that there exist mutually exclusive point neighborhoods U and V of P_1 and P_2 such that $R^*_{UV} \subset D$. If W is a neighborhood of Q within D', then U, V, and W are non-collinear. Suppose that, for each neighborhood U' of P_1 within U and each neighborhood V' of P_2 within V, $R^*_{U'V'}$ contains a plane. Then Condition (3) of Definition 4a cannot be met since $R^*_{U'V'}\cdot W=0$ and, by Condition (4) of Definition 4a, $R^*_{U'V'} \subset R^*_{U'V'W'}$, where W' is any neighborhood of Q within W. It follows from this remark that every plane contains more than one line.

EXAMPLE 1a. In the Euclidean space E_4, let the word line have its ordinary significance, but let "line" denote a plane of Example 1. Since every line is contained in some "line," the set of all planes in a bounded convex portion of E_4 is both line continuous and "line" continuous. But this set of planes fails to meet Condition (4) of Definition 4a with respect to "lines."

EXAMPLE 2. Let A, B, and C be non-collinear points in the space S. Let the flat π_{ABC} be defined as the point set sum of all lines l which contain a "vertex" and a point of the "opposite side" of the "triangle" ABC. If it is postulated that a line is contained in a flat if the line contains two distinct points of the flat, then a flat is a point continuum and the set of all flats is line continuous. To show this, let π_{ABC} be a flat and let A', B', and C' be non-collinear points of π_{ABC}. Then $\pi_{A'B'C'}=\pi_{ABC}$. (If P is a point of $\pi_{A'B'C'}$, and if A', B', and C' are in π_{ABC}, then P is in π_{ABC}. Hence $\pi_{A'B'C'} \subset \pi_{ABC}$. Thus, if C' is in $\overline{BC}$, then C is in $\overline{BC'}$ and $\pi_{ABC}=\pi_{ABC'}$. The same result follows when C' is in $\overline{AC}$. If C' is in π_{ABC} but not in $\overline{AC}$ or $\overline{BC}$, then either $\overline{C'A}$ meets $\overline{BC}$ in a point D, or $\overline{C'B}$ meets $\overline{AC}$, or $\overline{C'C}$ meets $\overline{AB}$. In the first case, $\pi_{ABC}=\pi_{ABD}=\pi_{ABC'}$, the other cases being similar. It follows that $\pi_{ABC}=\pi_{ABC'}=\pi_{AB'C'}=\pi_{A'B'C'}$.) Let $\{\pi_n\}$ be a sequence of flats containing three sequences $\{A_n\}$, $\{A_n'\}$, and $\{A_n''\}$ of points with non-collinear sequential limits A, A', and

A'' in a flat π, where A_n, A_n', and A_n'' are in π_n for each n. It may be assumed that, for each n, A_n, A_n', and A_n'' are non-collinear. Hence $\pi_n = \pi_{A_n A_n' A_n''}$. Let $\{B_n\}$ be a sequence of points with B_n in π_n for each n. For infinitely many n, either $\overline{A_n B_n}$ or $\overline{A_n' B_n}$ or $\overline{A_n'' B_n}$ has a point D_n in common with the respective line $\overline{A_n' A_n''}$ or $\overline{A_n A_n''}$ or $\overline{A_n A_n'}$; say it is $\overline{A_n B_n}$. Since lines are point continuous, some subsequence of $\{D_n\}$ has a sequential limit D in $\overline{A'A''}$. Hence some subsequence of $\{B_n\}$ has a sequential limit B in $\overline{AD}$ and therefore in $\pi_{AA'A''}$. By Remark 2, Condition (1) of Definition 2 is met. The proof of Condition (2) is similar. Thus the set of all flats is line continuous; by Lemma 2, every flat is closed. Since every flat is connected, it is a continuum. It follows that if Condition (3) of Definition 4a is postulated for flats, then flats are planes.

LEMMA 3a. *If U, V, and W are non-collinear open point sets, then the set of planes R_{UVW} is a plane neighborhood.*

The proof of this lemma is effected in three stages. Let A and B be points which are non-collinear with the open set W. An argument similar to that used in the proof of Lemma 3 shows that the "sheaf" σ of all planes containing $\overline{AB}$ and a point of W is such that $\sigma^* - \overline{AB}^*$ is open. Now let A be a point and let V and W be open sets such that A, V, and W are non-collinear. An argument similar to the second argument used in the proof of Lemma 3 shows that the "bundle" β of all planes containing A, a point of V, and a point of W is such that $\beta^* - A$ is open. A repetition of this argument, using this last result, leads to the lemma.

LEMMA 4. *If l is a line, if D is an open point set containing l^*, and if P and Q are distinct points of l, then there exists a pair of mutually exclusive point neighborhoods U of P and V of Q such that $R_{UV}^* \subset D$.*

Suppose there does not exist such a pair of neighborhoods. Let $\{U_n\}$ and $\{V_n\}$ be sequences of point neighborhoods which close down on† P and Q. It may be assumed that, for each n, U_n and V_n are mutually exclusive. For each n, there exists a point Z_n of $r_{U_n V_n}^*$ not in D. Hence there exists, for each n, a line of $r_{U_n V_n}$ containing Z_n, a point X_n of U_n, and a point Y_n of V_n. Since $\lim X_n = P$ and $\lim Y_n = Q$, it follows by Definition 2 that some subsequence of $\{Z_n\}$ has a sequential limit Z in l. But D is open and contains l. Hence the original supposition is inadmissible.

LEMMA 4a. *If p is a plane, if D is an open point set containing p^*, and if P, Q, and R are non-collinear points of p, then there exist non-collinear point neighborhoods U, V, and W of P, Q, and R respectively such that $R_{UVW}^* \subset D$.*

† That is, $\Pi U_n = P$, $U_n \supset \overline{U}_{n+1}$, and if T is any point neighborhood of P, $T \supset \overline{U}_n$ for some n.

It was pointed out in Remark 4a that there exist non-collinear neighborhoods U, V, and W of P, Q, and R. The rest of the proof is similar to the proof of Lemma 4.

LEMMA 5. *If l and l_1 are distinct lines, there exists an open point set containing l^* but not l_1^*. If p and p_1 are distinct planes, there exists an open set containing p^* but not p_1^*.*

The lemma follows immediately from Remarks 4 and 4a, the fact that two lines have at most one common point and two planes have at most one common line, and the fact that lines and planes are closed.

DEFINITION 5. *If H is a collection of lines and if P and Q are distinct points, then P and Q are called H-related if, for arbitrary point neighborhoods U of P and V of Q, there exists a line $l \neq \overline{PQ}$ of H which contains a point of U and a point of V. If H is a collection of planes and if P, Q, and R are non-collinear points, then P, Q, and R are called H-related if, for arbitrary point neighborhoods U, V, and W of P, Q, and R, there exists a plane $p \neq \overline{PQR}$ of H which contains points of U, V, and W.*

LEMMA 6. *In order that a line l be a limit line of a collection H of lines it is necessary and sufficient that l contain at least one pair of H-related points.*

Necessity. Suppose no two distinct points of l are H-related. Let P and Q be any two distinct points of l. By hypothesis, there exist point neighborhoods U of P and V of Q such that no line of H, except possibly $\overline{PQ}$, contains a point of U and a point of V. It may evidently be assumed that U and V are mutually exclusive. By Condition (3) of Definition 4, and Lemma 3, there exists a line neighborhood $r_{U'V'}$ of l which contains no line of H except possibly $\overline{PQ}$. The original supposition is thus inadmissible.

Sufficiency. Suppose that P and Q are H-related points of l. Let R be an arbitrary line neighborhood of l. By Lemma 4, there exists a pair of mutually exclusive point neighborhoods U of P and V of Q such that $R_{UV} \subset R$. Since P and Q are H-related, R_{UV}, and hence also R, contain a line of H other than $\overline{PQ}$.

LEMMA 6a. *In order that a plane p be a limit plane of a collection H of planes it is necessary and sufficient that p contain at least one set of three non-collinear H-related points.*

The proof is similar to the proof of Lemma 6.

LEMMA 7. *If R is a line neighborhood, then $\overline{R}^* \subset \overline{R^*}$.*

Suppose there exists a point A of $\overline{R}^*$ which is not contained in $\overline{R^*}$. Let l be a line of $\overline{R}$ containing A. Then l is a limit line of R. By Lemma 6, l contains

at least one pair of R-related points P and Q. Let $\{U_n\}$ and $\{V_n\}$ be sequences of point neighborhoods which close down on P and Q respectively. For each n, there exists a line l_n of R containing a point P_n of U_n and a point Q_n of V_n. Since $\lim P_n = P$ and $\lim Q_n = Q$, it follows that l and the sequence $\{l_n\}$ are in the relation described in Definition 2; but it is apparent that Condition (2) of Definition 2 cannot be met with regard to A.

LEMMA 7a. *If R is a plane neighborhood, then $\overline{R}^* \subset \overline{R^*}$.*

The proof is similar to the proof of Lemma 7.

THEOREM 1. *There exists a sequence $\{G_n'\}$ of collections of line neighborhoods satisfying Conditions* (1), (2), *and* (3) *of Axiom* 1 *when S is interpreted to be $G^{(1)}$ and "point" is interpreted to be "line."*

For each integer n and each line l let D_{nl} be the sum of all the neighborhoods of G_n which contain at least one point of l, and let G_{nl}' be the collection of all line neighborhoods R such that $R^* \subset D_{nl}$. For each integer n let G_n' be the sum of all the collections G_{nl}'. It is evident that the sequence $\{G_n'\}$ satisfies Conditions (1) and (2) of Axiom 1. To show that it satisfies Condition (3), let R be an arbitrary line neighborhood, and let x and y be lines of R, distinct or not. It is sufficient to consider the case $x \neq y$. By Lemma 5, there exists an open point set D containing x^* but not y^*. D may obviously be assumed to lie within R^*. By Lemma 4, there exists a line neighborhood R' containing x such that $R'^* \subset D$.

Suppose there exists no integer m such that, if g is a line neighborhood of G_m' containing x, then $\bar{g}$ is necessarily a subset of R'. In this event there exists a sequence $\{g_n\}$ of line neighborhoods, where, for each n, g_n belongs to G_n', contains x, and $\bar{g}_n$ is not contained in R'. Let c_n be a line such that $g_n^* \subset D_{nc_n}$. Since x is in g_n, $x^* \subset D_{nc_n}$. Let P and Q be distinct points of x. Because D_{nc_n} is a proper covering of c_n, there exist neighborhoods U_n and V_n of G_n containing P and Q respectively and also containing points P_n and Q_n respectively of c_n. It follows by (3) of Axiom 1 that $\lim P_n = P$ and $\lim Q_n = Q$. Hence, by an argument similar to that used in the proof of Lemma 4, if E is an open set containing x^*, there exists an integer m such that $c_n^* \subset E$ for all $n \geqq m$.

Since x is compact (Lemma 2), there exists† an open set E containing x such that $\overline{E} \subset R'^*$. Let m_0 be such an integer that, for all $n \geqq m_0$, $c_n^* \subset E$. The point set $x^* + c_{m_0}^* + c_{m_0+1}^* + \cdots$ is closed, compact, and contained in E. By Condition (4) of Axiom 1, there exists an integer $m \geqq m_0$ such that $D_{mx} + D_{mc_{m_0}} + D_{mc_{m_0+1}} + \cdots$ is contained in E. By Lemma 7, $(\bar{g}_m)^* \subset \overline{E}$. Thus the original supposition is inadmissible and the proof is complete.

† M, Chapter I, Theorem 23.

THEOREM 1a. *There exists a sequence* $\{G_n''\}$ *of collections of plane neighborhoods satisfying Conditions* (1), (2), *and* (3) *of Axiom* 1 *when* S *is interpreted to be* $G^{(2)}$ *and "point" is interpreted to be "plane."*

The proof is similar to the proof of Theorem 1.

Theorems 1 and 1a enable† us to quote theorems from Chapter I of M with regard to collections of lines in the line space $G^{(1)}$ and with regard to collections of planes in the plane space $G^{(2)}$.

DEFINITION 6. *Let* K *be a point set and let* P *be a limit point of* K. *A line* l *is called a tangent line of* K *at* P *if, for an arbitrary point neighborhood* V *of* P *and an arbitrary line neighborhood* R *of* l, *there exists a line of* R *containing at least two distinct points of* KV. *The set of all tangent lines of* K *at* P *is called the paratingent of* K *at* P.

DEFINITION 6a. *Let* K *be a point set and let* P *be a limit point of* K. *A plane* p *is called a tangent plane of* K *at* P *if, for an arbitrary point neighborhood* V *of* P *and an arbitrary plane neighborhood* R *of* p, *there exists a plane of* R *containing at least three non-collinear points of* KV. *The set of all tangent planes of* K *at* P *is called the biparatingent of* K *at* P.

REMARK 5. For a line l to be a tangent line of a point set K at a limit point P of K it is necessary that l contain P. It is evident that K has no tangent line at a point which is not a limit point of K. Similar comments apply to tangent planes.

THEOREM 2. *The paratingent* Λ *of a point set* K *at a limit point* P *of* K *is closed.*

If l is a limit line of Λ, and if R is a line neighborhood of l, then R contains a line l_1 of Λ. Since l_1 is in Λ, there exists, for an arbitrary point neighborhood V of P, a line l' of R which contains two distinct points of KV. Thus l is in Λ.

THEOREM 2a. *The biparatingent* Π *of a point set* K *at a limit point* P *of* K *is closed.*

The proof is similar to the proof of Theorem 2.

REMARK 6. Definition 6 is not essentially altered if the clause "there exists a line of R containing at least two distinct points of KV" is replaced by the clause "there exists a line of R which contains at least two distinct points of KV or is a limit line of lines containing at least two distinct points of KV." This remark may be substantiated by an argument similar to that used in the proof of Theorem 2. A similar remark applies to Definition 6a.

† M, p. 82.

Lemma 8. *If H is a collection of lines containing a common point P, then a necessary condition that H be the sum of two mutually separated*† *collections M and N is that P be a cut point of the connected point set $H^*=M^*+N^*$.*

Suppose that H^*-P is connected. Then at least one of the sets M^*-P and N^*-P, say the former, contains a limit point Q of the other. It is evident that P and Q are N-related. By Lemma 6, $\overline{PQ}$ is a limit line of N. But $\overline{PQ}$ is a line of M. This contradicts the hypothesis that M and N are mutually separated. Hence H^*-P is not connected and P is a cut point of the point set H^*.

Lemma 8a. *If H is a collection of planes containing a common line l, then a necessary condition that H be the sum of two mutually separated collections M and N is that l be a cut line of the connected line set $H^{**}=M^{**}+N^{**}$, where H^{**} denotes the collection of all lines contained in any plane of H.*

It follows from Lemma 9 below and the fact that a line is a continuum that if A is a point and if l is a line not containing A, the set of all lines containing A and a point of l is connected. Hence it follows that a plane is a connected set of lines. (Let p be a plane and let l and l' be distinct lines of p. Let A and B be points of l and let A' and B' be points of l', where all four of these points are distinct. The set of all lines containing A and a point of $\overline{A'B}$ is connected; this set contains l and $\overline{AA'}$. The set of all lines containing A' and a point of $\overline{AB'}$ is connected; this set contains l' and $\overline{AA'}$. Hence the sum of these two sets is connected and contains l and l'.) It therefore follows that H^{**} is connected.

Now suppose that $H^{**}-l$ is connected. Then at least one of the sets $M^{**}-l$ and $N^{**}-l$, say the former, contains a limit line l' of the other. The line l' is coplanar with l and determines a plane $\overline{ll'}$. If P is a point of l' not in l, then, by Definition 2, P is a limit point of the point set N^*-l. If Q and R are distinct points of l not in l', then P, Q, and R are N-related. The remainder of the proof follows the line of argument in the preceding proof.

Axiom 2. *The space S is compact.*

Theorem 3. *The line space $G^{(1)}$ is compact.*

Let $\{l_n\}$ be a sequence of distinct lines and let $\{P_n\}$ be a sequence of points with P_n in l_n for each n. By Axiom 2, some subsequence $\{P_{n_i}\}$ of $\{P_n\}$ has a sequential limit P. By Remark 4, there exists a neighborhood V of P which contains no line. Let $\{Q_{n_i}\}$ be a sequence of points where, for each i, Q_{n_i} is in l_{n_i} and not in V. By Axiom 2, some subsequence of $\{Q_{n_i}\}$ has a

† Two sets are said to be mutually separated if they are mutually exclusive and if neither contains a limit element of the other.

sequential limit Q, and $Q \neq P$. By Lemma 6, some subsequence of $\{l_n\}$ has a sequential limit. An obvious extension of this argument proves

THEOREM 3a. *The plane space $G^{(2)}$ is compact.*

Notation. If K is a point set and if A is a point, then $K;A$ denotes the closure of the set of all lines containing A and a point of $K-KA$. If B and B' are points of K distinct from A, and if there exists a connected subset of $K;A$ containing $\overline{BA}$ and $\overline{AB'}$, then $K;B,A,B'$ denotes the component of $K;A$ containing $\overline{BA}$ and $\overline{AB'}$. If k is a set of lines with a common point A and if a is a line containing A, then $k;a$ denotes the closure of the set of all planes containing a and a line of $k-ka$. If b and b' are lines of k distinct from a, and if there exists a connected subset of $k;a$ containing $\overline{ba}$ and $\overline{ab'}$ (where $\overline{ba}$ is the plane containing b and a), then $k;b,a,b'$ denotes the component of $k;a$ containing $\overline{ba}$ and $\overline{ab'}$.

LEMMA 9. *If K is a connected point set, if A is a point, and if B and B' are points of K distinct from A, then either $K;B,A,B'$ exists or else $K;A,B,B'$ and $K;B,B',A$ exist.*

It may be assumed that A, B, and B' are non-collinear, as otherwise the lemma is trivial. Suppose there exists no connected subset of $K;A$ containing $\overline{BA}$ and $\overline{AB'}$. Since $K;A$ is closed and compact, $K;A$ consists† of the sum of two mutually separated line sets M and N containing $\overline{BA}$ and $\overline{AB'}$ respectively. By Lemma 8 it follows that M^*-A and N^*-A, and hence $M^*(K-KA)$ and $N^*(K-KA)$, are mutually separated. Hence A is a point of K and A separates B and B' in K. By a well known theorem,‡ B does not separate A and B' in K, and B' does not separate A and B in K. Therefore $K;A,B,B'$ and $K;B,B',A$ exist.

LEMMA 9a. *If K is a connected set of lines with a common point A, if a is a line containing A, and if b and b' are lines of k distinct from a, then either $k;b,a,b'$ exists or else $k;a,b,b'$ and $k;b,b',a$ exist.*

The proof is exactly parallel to the proof of Lemma 9.

Notation. If K is a connected point set, if A is a point, and if B and B' are points of K distinct from A, then $K;B-A-B'$ denotes $K;B,A,B'$ if this set exists and $K;A,B,B'+K;B,B',A$ otherwise. Since $K;A,B,B'$ and $K;B,B',A$ have a common line, it follows by Lemma 9 that $K;B-A-B'$ always exists, contains $\overline{BA}$ and $\overline{AB'}$, is connected, and each line of $K;B-A-B'$ contains at least two distinct points of the point set $K+A$ or is a limit line of such lines.

† M, Chapter I, Theorem 35.

‡ M, Chapter I, Theorem 49.

LEMMA 10. *If $\{M_n\}$ is a sequence of connected sets of lines (planes), and if there exists a convergent sequence $\{m_n\}$ of lines (planes) with m_n in M_n for each n, then the limiting set of $\{M_n\}$ is a line (plane) continuum.*

This lemma occurs in M, Chapter I, as Theorem 42.

THEOREM 4. *If K is either a locally connected point set or a point continuum, and if P is a limit point of K, then the paratingent Λ of K at P is a continuum of lines.**

The proof of this theorem is carried out first for the case where K is a continuum. Let $\{V_n\}$ be a sequence of point neighborhoods which closes down on P. Let ρ and σ be distinct lines of Λ (assuming Λ to contain more than one line). Let $\{R_n\}$ and $\{S_n\}$ be sequences of line neighborhoods closing down on ρ and σ respectively. By Definition 6, there exists a pair of distinct points A_n and B_n in KV_n such that the line $\overline{A_nB_n}$ is in R_n; likewise there exists a pair of distinct points C_n and D_n in KV_n such that the line $\overline{C_nD_n}$ is in S_n. Of the two points C_n and D_n, let C_n be taken as one which is distinct from B_n. Let it be supposed that V_1 was so chosen that K contains a point outside V_1. For $n>1$, let β_n (γ_n) be that component of $K\overline{V}_2$ containing B_n (C_n). Since K is a continuum, there exists† a point B_n' (C_n') of β_n (γ_n) on the boundary of V_2. Since the boundary of V_2 is compact, there exists a subsequence $\{n_i\}$ of the sequence $\{n\}$ such that the sequence $\{B_{n_i}'\}$ has a sequential limit B'; likewise, there exists a subsequence $\{n_{i_j}\}$ of $\{n_i\}$ such that the sequence $\{C_\alpha\}$ has a sequential limit C', where α runs (both now and in what follows) through the sequence $\{n_{i_j}\}$. By Lemma 9, there exist, for each α, connected line sets $\beta_\alpha;B_\alpha-A_\alpha-B_\alpha'$, $\beta_\alpha;B_\alpha-C_\alpha-B_\alpha'$, $\gamma_\alpha;C_\alpha-B_\alpha-C_\alpha'$, and $\gamma_\alpha;C_\alpha-D_\alpha-C_\alpha'$, where each line of each of these four sets contains two distinct points of KV_1 or is a limit line of such lines. Since $\beta_\alpha;B_\alpha-C_\alpha-B_\alpha'$ and $\gamma_\alpha;C_\alpha-B_\alpha-C_\alpha'$ have the line $\overline{B_\alpha C_\alpha}$ in common, $\beta_\alpha;B_\alpha-C_\alpha-B_\alpha'+\gamma_\alpha;C_\alpha-B_\alpha-C_\alpha'$ is connected. But $\lim \overline{B_\alpha A_\alpha}=\rho$ and $\lim \overline{A_\alpha B_\alpha'}=\overline{PB'}$. Hence ρ and $\overline{PB'}$ are lines of the limiting set Σ^1 of the sequence $\{\beta_\alpha;B_\alpha-A_\alpha-B_\alpha'\}$. Similarly, $\overline{PB'}$ and $\overline{PC'}$ are lines of the limiting set Σ^2 of $\{\beta_\alpha;B_\alpha-C_\alpha-B_\alpha'+\gamma_\alpha;C_\alpha-B_\alpha-C_\alpha'\}$, and $\overline{PC'}$ and σ are lines of the limiting set Σ^3 of $\{\gamma_\alpha;C_\alpha-D_\alpha-C_\alpha'\}$. By Lemma 10, Σ^1, Σ^2, and Σ^3 are line continua; because pairs of these sets have common lines, their sum $\lambda_1=\Sigma^1+\Sigma^2+\Sigma^3$ is a line continuum containing ρ and σ. Every line of λ_1 contains at least two distinct points of KV_1 or is a limit of such lines. Let λ_n be the line continuum obtained from KV_n by the same process that led to λ_1 from KV_1. By Lemma 10, the

* The central idea underlying the proof of this theorem was suggested by a proof of Mirguet, Annales de l'École Normale Supérieure, vol. 51 (1934), p. 201.

† M, Chapter I, Theorem 40.

limiting set λ of the sequence $\{\lambda_n\}$ is a line continuum; it is evident that λ contains ρ and σ. But by Remark 6, each line of λ is a tangent line of K at P. Hence Λ is connected. By Theorem 2, Λ is closed. Therefore Λ is a continuum of lines.

The proof of the theorem in the case where K is locally connected is similar to the above proof. Within each neighborhood V_n of the above sequence $\{V_n\}$ there exists a neighborhood V_n' of P such that KV_n' is connected. The points A_n, B_n, C_n, and D_n may be taken in KV_n'. By Lemma 9, the sets $KV_n';A_n-B_n-C_n$ and $KV_n';B_n-C_n-D_n$ exist. Since these sets have the common line $\overline{B_nC_n}$, their sum is connected. The limit process based on Lemma 10 may be applied as before to these sets.

Notation. If K is a point set, if A and B are distinct points, and if C and C' are points of K not in $\overline{AB}$, then $(K;A);\overline{AC},\overline{AB},\overline{AC'}$, if it exists, will be denoted by

$$K;\,\underset{B}{C,A,C'}.$$

If K is connected then, by Lemma 9, either

(1) $K;C,A,C'$ exists, or

(2) $K;A,C,C' + K;C,C',A$ exists.

In Case (1) it follows by Lemma 9a that either

(1a) $K;\,\underset{B}{C,A,C'}$ exists, or

(1b) $K;\,\underset{C}{B,A,C'} + K;\,\underset{C'}{C,A,B}$ exists.

In Case (2) either

(2a) $K;\,\underset{B}{A,C,C'} + K;\,\underset{B}{C,C',A}$ exists, or

(2b) $K;\,\underset{A}{B,C,C'} + K;\,\underset{C'}{A,C,B} + K;\,\underset{B}{C,C',A}$ exists, or

(2c) $K;\,\underset{B}{A,C,C'} + K;\,\underset{C}{B,C',A} + K;\,\underset{A}{C,C',B}$ exists, or

(2d) $K;\,\underset{A}{B,C,C'} + K;\,\underset{C'}{A,C,B} + K;\,\underset{C}{B,C',A} + K;\,\underset{A}{C,C',B}$ exists.

Of the six sets (1a) to (2d), let

$$K;\, C-\frac{A}{B}-C'$$

denote the first one existing. If C and C' are collinear with A or B, then

$$K;\, C-\frac{A}{B}-C'$$

denotes the single plane $\overline{ABC}=\overline{ABC'}$. Thus, for any connected point set K and for points A, B, C, and C' as described above,

$$K;\, C-\frac{A}{B}-C'$$

always exists, contains $\overline{CAB}$ and $\overline{ABC'}$, is connected, and each plane of

$$K;\, C-\frac{A}{B}-C'$$

contains at least three non-collinear points of $K+A+B$ or is a limit plane of such planes.

The extension of Theorem 4 to tangent planes in the case where K is a point continuum seems to present difficulty in that the non-collinear point triples needed for the proof may not exist. However, the following lemma makes it possible to prove the extension in two special cases of interest.

LEMMA 11. *If P is a point, if $\{\alpha_n\}$ and $\{\beta_n\}$ are two sequences of point sets, if the sequential limiting set of $\{\alpha_n\}$ contains a point A distinct from P, if the limiting set of every subsequence of $\{\beta_n\}$ contains a point distinct from P, and if it is impossible to find a subsequence $\{n_i\}$ of $\{n\}$ and sequences $\{A_{n_i}\}$ and $\{B_{n_i}\}$ such that* $\lim \overline{A_{n_i}B_{n_i}}$ *exists and does not contain P, where, for each n_i, A_{n_i} is a point of α_{n_i}, B_{n_i} is a point of β_{n_i}, and $A_{n_i}\neq B_{n_i}$, then the limiting sets of $\{\alpha_n\}$ and $\{\beta_n\}$ are contained in the line $\overline{PA}$.*

The proof of this lemma is readily effected by first showing that an arbitrary point of the limiting set of $\{\beta_n\}$ is contained in $\overline{PA}$ and then using this result to show that an arbitrary point of the limiting set of $\{\alpha_n\}$ is contained in $\overline{PA}$.

Let K be a point continuum and let P be a point of K. Let $\{V_n\}$ be a sequence of point neighborhoods closing down on P. Let ρ and σ be distinct planes of the biparatingent Π of K at P. Let $\{R_n\}$ and $\{S_n\}$ be sequences of plane neighborhoods closing down on ρ and σ respectively. By Definition 6a, there exist three non-collinear points A_n, B_n, and C_n in KV_n such that the plane $\overline{A_nB_nC_n}$ is in R_n; likewise there exist three non-collinear points X_n, Y_n, and Z_n in KV_n such that the plane $\overline{X_nY_nZ_n}$ is in S_n. Let α_n, β_n, γ_n, ξ_n, η_n, ζ_n be those components of $K\overline{V}_2$ containing A_n, B_n, C_n, X_n, Y_n, and Z_n.

If no non-degenerate subcontinuum of that part of K within some neighborhood V of P is contained in a line, and if V_1 is taken within V, then by Lemma 11 it is possible to find a subsequence $\{n_i\}$ of $\{n\}$ and points B_{n_i}', C_{n_i}', X_{n_i}', and Y_{n_i}' of β_{n_i}, γ_{n_i}, ξ_{n_i}, and η_{n_i} such that $\lim \overline{B_{n_i}'C_{n_i}'}$, $\lim \overline{C_{n_i}'X_{n_i}'}$, and $\lim \overline{X_{n_i}'Y_{n_i}'}$ do not contain P. Hence there exists a set of connected sets of planes of the following sort for sufficiently large n_i:

$$(1)\qquad \begin{aligned}
&\gamma_{n_i};\ C_{n_i} \overset{A_{n_i}}{\underset{B'_{n_i}}{-}} C'_{n_i} + \beta_{n_i};\ B_{n_i} \overset{A_{n_i}}{\underset{C'_{n_i}}{-}} B'_{n_i},\\
&\beta_{n_i};\ B'_{n_i} \overset{X_{n_i}}{\underset{C'_{n_i}}{-}} B_{n_i} + \xi_{n_i};\ X_{n_i} \overset{B_{n_i}}{\underset{C'_{n_i}}{-}} X'_{n_i},\\
&\gamma_{n_i};\ C'_{n_i} \overset{Y_{n_i}}{\underset{X'_{n_i}}{-}} C_{n_i} + \eta_{n_i};\ Y_{n_i} \overset{C_{n_i}}{\underset{X'_{n_i}}{-}} Y'_{n_i},\\
&\eta_{n_i};\ Y'_{n_i} \overset{X'_{n_i}}{\underset{Z_{n_i}}{-}} Y_{n_i} + \xi_{n_i};\ X'_{n_i} \overset{Y_{n_i}}{\underset{Z_{n_i}}{-}} X_{n_i}.
\end{aligned}$$

If the preceding hypothesis concerning the subcontinua of K is not valid, it is still possible to obtain a set of planes of the form (1) by imposing conditions on the lines and planes of the space S like the following: (a) if P is a point of the space S, there exists a point neighborhood V of P such that, if π is a plane containing a point of V, then $V - V\pi$ is the sum of two mutually separated point sets; and (b) a line is a simple continuous arc. Having obtained the set (1), one may prove the extension of Theorem 4 by following the line of argument in the proof of Theorem 4.

THEOREM 4a. *If K is a locally connected point set and if P is a limit point of K, then the biparatingent Π of K at P is a continuum of planes.*

The proof of this theorem is effected by taking neighborhoods V_n' of P within V_n such that KV_n' is connected and by modifying the above argument in exactly the same way as in the last paragraph of the proof of Theorem 4.

It is readily possible to define hyperplanes of every finite "dimension" with the aid of Definition 2 and the inductive process indicated in Definition 4a. Most of the results of this paper may be extended by means of these hyperplanes.

CASE SCHOOL OF APPLIED SCIENCE,
CLEVELAND, OHIO

ON RINGS OF OPERATORS. II*

BY

F. J. MURRAY† AND J. VON NEUMANN

Introduction. This paper is a continuation of one by the same authors: *On rings of operators*, Annals of Mathematics, (2), vol. 37 (1936), pp. 116–229. It contains the solution of certain problems which were left open there. We will prove the general additivity of trace $Tr_M(A)$, its weak continuity, and certain isomorphisms between $\mathfrak{H}$, M, and M' (cf. the remarks (i)–(iv) at the end of the above quoted paper). All these considerations refer to "Case (II)" for M (cf. Theorem VIII, loc. cit.).

The properties of $Tr_M(A)$ are established by obtaining for it a representation

$$Tr_M(A) = \sum_{i=1}^{m} (Ag_i, g_i)$$

(with a fixed, finite $m=1, 2, \cdots$, and fixed $g_1, \cdots, g_m \epsilon \mathfrak{H}$). This representation is remarkable, because it is obviously a close analogue of the representation of $Tr_M(A)$ as a trace, that is, as the arithmetic mean of the diagonal matrix-elements of A in the cases (I_n), $n=1, 2, \cdots$, when M is essentially the full matrix ring of an n-(finite-) dimensional Euclidean space.

For certain cases (with the help of which the others are then mastered) we have even $m=1$.

In Part I the above representation of $Tr_M(A)$ is obtained approximately. The technically interested reader may find it worth observing that the exhaustion method we use there (§§1.2 and 1.3) is analogous to certain procedures which can be used advantageously in the theories of measures and integration too. On this basis we establish the main properties of $Tr_M(A)$ in Part II, and then obtain the exact representation of $Tr_M(A)$ in Part III. Here two maximum-problems, called (A) and (B), which seem to possess some independent interest too, play a decisive role.

Part IV is devoted to establishing an isomorphism between $\mathfrak{H}$, M, and M'. It turns out that a certain algebraic-topological extension $Q(M)$ of M is isomorphic to $\mathfrak{H}$ and that M and M' play in it the role of right- and left-multiplication. This leads to an interesting and entirely new type of infinite hypercomplex systems, which are at the same time Hilbert spaces. A subsequent paper will be devoted to their independent study.

* Presented to the Society, October 31, 1936; received by the editors July 28, 1936.

† National Research Fellow.

The appendix deals with the possibility of considering $\boldsymbol{M}$ (in the case (II_1)) as a system of matrices with continuously spread rows and columns.

We will use the notations, definitions, and results of our paper *On rings of operators*, quoted above, throughout this paper. We will quote it, whenever necessary, as R.O. All other quotations follow the bibliography of R.O. (pp. 125–126, Nos. (1)–(22)).

The isomorphism problems of different rings $\boldsymbol{M}$ of class (II_1) (cf. the remark (v) at the end of R.O.) are not discussed here. They will be dealt with in a subsequent publication.

Since the appearance of R.O. the second-named author has succeeded in finding new representations of case (II_1) in terms of infinite direct products, which throw new light on (II_1) as a limiting case of the (I_n), $n=1, 2, \cdots$, as well as a way of applying the present theory to quantum-mechanics. These subjects too will be discussed in papers which will follow soon.

CONTENTS

Chapter I. Approximate form of $Tr_M(A)$ (for $\alpha \geqq 1$)

1.1. We assume that we have given a factor $\boldsymbol{M}$ in case II_1. Now we normalize D_M and $D_{M'}$ so that $C=1$ (cf. R.O., pp. 179–182). This means (R.O., loc. cit.) that for every $f \epsilon \mathfrak{H}$, $D_M(\mathfrak{M}_f^{M'}) = D_{M'}(\mathfrak{M}_f^{M})$ and that the range of D_M is the interval $0 \leqq x \leqq 1$. Let the range of $D_{M'}$ be the interval $0 \leqq x \leqq \alpha$, where $0 < \alpha \leqq \infty$. In this part, we assume that $\alpha \geqq 1$.

Now $\boldsymbol{M}'$ is a factor of class II_1 or II_∞ by R.O., Theorem X. In the first case the range of $D_{M'}$ is in the normalization (loc. cit.) the interval, $0 \leqq x \leqq 1$, and $C = 1/\alpha$ (since $D_{M'} = 1/\alpha$ times its value in the previous paragraph). Since $C \leqq 1$, we have in both cases by the discussion on page 182 of R.O., that $\Delta_0 = \Delta$. By R.O., Definition 10.1.1, this implies that there exists an f such that $D_M(\mathfrak{M}_f^{M'}) = 1$. Since we can pass from the normalization of this paragraph to that of the previous paragraph by multiplying $D_{M'}$ by α and leaving D_M unchanged, we see that this holds even with our previous normalization.

Thus we have an f such that $D_M(E_f^{M'}) = 1 = D_M(1)$ or $D_M(1 - E_f^{M'}) = 0$ which implies $1 - E_f^{M'} = 0$; that is, $E_f^{M'} = 1$. We now assume that such an f has been chosen and is held fixed for the following discussion.

1.2. With respect to this fixed f, we now define certain relations concerning a projection $E \epsilon \boldsymbol{M}$.

Definition 1.2.1. *Let $E \neq 0$ be a projection, $\epsilon \boldsymbol{M}$, θ a real number $\geqq 0$. The relation $E > \theta$ ($E \geqq \theta$, $E < \theta$, $E \leqq \theta$) is said to hold if $(Ef, f) > \theta D_M(E)$ ($(Ef, f) \geqq \theta D_M(E)$, $(Ef, f) < \theta D_M(E)$, $(Ef, f) \leqq \theta D_M(E)$ respectively).*

The relation $E \geqq_p \lambda$ ($E \leqq_p \lambda$) is said to hold for a projection $E \neq 0$ if for every $F \epsilon \boldsymbol{M}$ such that $F \leqq E$, $F \neq 0$, we have $F \geqq \lambda$ ($F \leqq \lambda$).

Lemma 1.2.1. *Let $\{E_i\}$ be a sequence of projections (finite or infinite) with $E_i \epsilon \boldsymbol{M}$, $E_i E_j = \delta_i^j E_i$. Then if for every i, $E_i \leqq \lambda$ ($E_i \geqq \lambda$), and if for some i, say i_0, $E_{i_0} < \lambda$ ($E_{i_0} > \lambda$) we have $\sum_i E_i < \lambda$ ($\sum_i E_i > \lambda$).*

We have, for every i, $\lambda D(E_i) \leqq (E_i f, f)$, hence $\lambda\sum_{i\neq i_0} D(E_i) \leqq \sum_{i\neq i_0}(E_i f, f)$ and $\lambda D(E_{i_0}) < (E_{i_0} f, f)$. These imply $\lambda\sum_i D(E_i) < \sum_i (E_i f, f)$ or $\lambda D(\sum_i E_i) < (\sum_i E_i f, f)$, which was to be shown.

Lemma 1.2.2. *If $E \geqq \lambda$ ($E \leqq \lambda$), then there exists an F such that $E \geqq F$, $F \geqq_p \lambda$ ($F \leqq_p \lambda$).*

By Definition 1.2.1 we have $E \neq 0$. Suppose there is no such F. Then there must be an E_1, such that $E \geqq E_1$, and $E_1 < \lambda$. (Otherwise E itself would be such an F.) Now let Ω be the first ordinal number which belongs to a cardinal number $\aleph > \aleph_0$. Let α be an ordinal number such that $\alpha < \Omega$. Let us suppose that for all $\beta < \alpha$ an E_β has been defined in such a way that $E_\beta \neq 0$, $E_\beta \leqq E$, $E_\beta < \lambda$, $E_{\beta_1} E_{\beta_2} = \delta_{\beta_1,\beta_2} E_{\beta_1}$. Now if $E - \sum_{\beta<\alpha} E_\beta = 0$, let $E_{\alpha'}$ be undefined for $\alpha' \geqq \alpha$. If, however, $E - \sum_{\beta<\alpha} E_\beta \neq 0$, by hypothesis, $E - \sum_{\beta<\alpha} E_\beta$ is not $\geqq_p \lambda$, hence there exists an E_0, with $E_0 \neq 0$, $E - \sum_{\beta<\alpha} E_\beta \geqq E_0$, $E_0 < \lambda$, $E \geqq E - \sum_{\beta<\alpha} E_\beta \geqq E_0$. Thus if we let $E_\alpha = E_0$, we have, if $E - \sum_{\beta<\alpha} E_\beta \neq 0$, an E_α which is orthogonal to all previous E_β and $E_\alpha < \lambda$.

Now $D(E) \geqq \sum_{\beta<\alpha} D(E_\beta)$ for every α. Since $D(E_\beta) > 0$, this implies that there is only denumerably many of the numbers $D(E_\beta)$. Hence for some $\alpha < \Omega$, $E - \sum_{\beta<\alpha} E_\beta = 0$. Inasmuch as $\alpha < \Omega$, we can re-index the E_β's into a finite or a simply infinite sequence. Now $E_1 < \lambda$, $E_i \leqq \lambda$, and since $E = \sum_i E_i$, Lemma 1.1 implies that $E < \lambda$, a contradiction.

Lemma 1.2.3. *If $E \geqq_p \lambda$ ($E \leqq_p \lambda$), then if $E \geqq F$, $F \neq 0$, $F \epsilon \boldsymbol{M}$, we have $F \geqq_p \lambda$ ($F \leqq_p \lambda$).*

If F_1 is such that $F \geqq F_1$, $F_1 \neq 0$ then $E \geqq F \geqq F_1$ and hence by Definition 1.2.1, $F_1 \geqq \lambda$. This statement implies that $F \geqq_p \lambda$.

1.3. Now $(1f, f) = \|f\|^2 = \|f\|^2 D(1)$. Now let $\theta_0 = \|f\|^2$ which is of course not zero for $\mathfrak{M}_f{}^M = \mathfrak{H}$ and this precludes $f = 0$. Thus 1 (the projection, not the number) is $\geqq \theta_0$. By Lemma 1.2.2, there exists an $E \epsilon \boldsymbol{M}$, such that $E \geqq_p \theta_0$. Now let λ be the least upper bound of the numbers θ such that $E \geqq_p \theta$. λ must be less than $(Ef, f)/D_M(E)$ and since $D_M(E) > 0$, it must be finite. Let ϵ be any number > 0. E is not $\geqq_p \lambda + \epsilon$. Hence there is an E_1, such that $E \geqq E_1$, and $E_1 \leqq \lambda + \epsilon$. Lemma 1.2.2 now implies that there is an E_2 with $E_1 \geqq E_2$ and $E_2 \leqq_p \lambda + \epsilon$.

Now if F is such that $E \geqq F$, $F \epsilon \boldsymbol{M}$, $F \neq 0$, we have $F \geqq \theta$ for all $\theta < \lambda$, which means of course that $F \geqq \lambda$ and $E \geqq_p \lambda$. Now since $E \geqq E_1 \geqq E_2$, Lemma 1.2.3 implies that $E_2 \geqq_p \lambda$. Thus for some fixed $\lambda > 0$, given any $\epsilon > 0$, we can find a non-zero $E_2 \epsilon \boldsymbol{M}$, such that for all $F \epsilon \boldsymbol{M}$ with the property $E_2 \geqq F$, we have

$$(\lambda + \epsilon) D_M(F) \geqq (Ff, f) \geqq \lambda D_M(F).$$

Now if we let f be $\lambda^{1/2} f$ above, $\epsilon = \lambda(K-1) E_2 = E$, we see that

LEMMA 1.3.1. *To every $K>1$ there exists an f and $E\epsilon M$, $E\neq 0$, such that for every $F\epsilon M$ with the property $E\geqq F$,*

$$KD_M(F) \geqq (Ff,f) \geqq D_M(F).$$

1.4. We can now show

LEMMA 1.4.1. *Let K, f, and E be as in Lemma 1.3.1. Let $A\epsilon M$ be a positive definite self-adjoint operator such that $EA=AE=A$. Then*

$$KTr_M(A) \geqq (Af,f) \geqq Tr_M(A).$$

Let c be the bound of A, $E(\lambda)$ the resolution of the identity corresponding to A. Now by R.O., pp. 212–213,

$$Tr_M(A) = \int_0^c \lambda dD(E(\lambda)); \qquad (Af,f) = \int_0^c \lambda d(E(\lambda)f,f).$$

Now since $AE=A$, $E(0)\geqq 1-E$, and since A commutes with E, E commutes with all $E(\lambda)$ (cf. (15), Theorem 8.2). Now $E(\lambda)=E(\lambda)E+E(\lambda)(1-E)$. Since $E(\lambda)$ and E commute, $E(\lambda)E=F(\lambda)$ is a projection and similarly $E(\lambda)(1-E)$ is a projection. For $\lambda\geqq 0$, we have $E(\lambda)(1-E)\leqq 1-E$ and also $E(\lambda)(1-E)\geqq E(0)(1-E)\geqq(1-E)^2=1-E$ and thus $E(\lambda)(1-E)=1-E$. So for $\lambda\geqq 0$, $E(\lambda)=F(\lambda)+(1-E)$. Now $F(\lambda)(1-E)=E(\lambda)E(1-E)=0$, hence by R.O., Definition 8.2.1, $D(E(\lambda))=D(F(\lambda))+D(1-E)$. Thus we have

$$Tr_M(A) = \int_0^c \lambda dD(E(\lambda)) = \int_0^c \lambda dD(F(\lambda)) + \int_0^c \lambda dD(1-E) = \int_0^c \lambda dD(F(\lambda))$$

and

$$(Af,f) = \int_0^c \lambda d(E(\lambda)f,f) = \int_0^c \lambda d(F(\lambda)f,f) + \int_0^c \lambda d((1-E)f,f)$$
$$= \int_0^c \lambda d(F(\lambda)f,f).$$

But Lemma 1.3.1 now implies that if $0\leqq\alpha<\beta\leqq c$, then

$$KD_M(F(\beta)-F(\alpha)) \geqq ((F(\beta)-F(\alpha))f,f) \geqq D_M(F(\beta)-F(\alpha))$$

which by the definition of the Riemann-Stieltjes integral yields that

$$KTr_M(A) \geqq (Af,f) \geqq Tr_M(A).$$

LEMMA 1.4.2. *For a projection $E\neq 0$, $E\geqq_p\theta$ ($\leqq_p\theta$) is equivalent to the statement that for all positive definite $A\epsilon M$ such that $EA=AE=A$,*

$$(Af,f) \geqq \theta Tr_M(A) \quad (\leqq \theta Tr_M(A)).$$

Since $Tr_M(F)=D_M(F)$, F a projection, the last statement implies the first. The converse is shown by a proof similar to that of Lemma 1.4.1.

LEMMA 1.4.3. *Let $A\epsilon M$ be such that $EA=AE=A$, then $\|A^*f\|^2\leqq K\|Af\|^2$, if E, f, K are as in Lemma* 1.3.1.

A^*A is positive definite and ϵM. Furthermore $AE=EA=A$ implies $EA^*=A^*E=A^*$, and thus $EA^*A=A^*A=A^*AE$. Hence Lemma 1.4.1 applies to A^*A and we have

$$(\alpha)\qquad KTr_M(A^*A) \geqq (A^*Af, f) \geqq Tr_M(A^*A).$$

Using the canonical decomposition (R.O., Definition 4.4.1) for A^* we have $A^*=UB$, where U may be taken as unitary. (In the finite cases, we see (cf. R.O., Lemma 16.1.1) that there exists a partially isometric V with initial set $(f;\, A^*f=0)$ and final set $(f;\, Af=0)$. Now if W is as in R.O., Definition 4.4.1, for $A=A^*$, let $U=W+V$.) B is self-adjoint and equals $(AA^*)^{1/2}$. Now $A=BU^*=BU^{-1}$ and hence $A^*A=UBBU^{-1}=UB^2U^{-1}=UAA^*U^{-1}$. Hence

$$(\beta)\qquad Tr_M(A^*A) = Tr_M(UAA^*U^{-1}) = Tr_M(AA^*).$$

Substituting A^* for A in our previous result we have

$$(\gamma)\qquad KTr_M(AA^*) \geqq (AA^*f, f) \geqq Tr_M(AA^*).$$

Combining (α), (β), and (ϕ), we obtain

$$K(A^*Af, f) \geqq KTr_M(A^*A) = K\,Tr_M(AA^*) \geqq (AA^*f, f).$$

Since $(A^*Af,f)=(Af, Af)=\|Af\|^2$, $(AA^*f,f)=(A^*f, A^*f)=\|A^*f\|^2$; this is the desired inequality.

1.5. Now if E is as in Lemma 1.3.1, let n be the smallest integer such that $1/n\leqq D_M(E)$. Then if $E^0\leqq E$ is such that $D_M(E^0)=1/n$, Lemma 1.3.1 will hold with E^0 in place of E. Now f was chosen in such a manner that $\mathfrak{M}_f^{M'}=\mathfrak{H}$. This implies that $\mathfrak{M}_{E^0f}^{M'}$ is the range of E^0, because since the set $(A'f;\, A'\epsilon M')$ is dense in $\mathfrak{H}$, the set $(E^0A'f;\, A'\epsilon M')=(A'E^0f;\, A'\epsilon M')$ must be dense in the range of E^0, or $E^0=E_{E^0f}^{M'}$. Since $C=1$, $D_{M'}(E_{E^0f}^{M})=D_M(E_{E^0f}^{M'})$ $=D_M(E^0)=1/n$.

Now let $E^0=E_1$, $E_{E^0f}^{M}=E_1'$. Since $D_M(E_1)=1/n$, $D_M(1)=1$, we can find n projections $\{E_j\}$, $j=1, \cdots, n$ with the first equal to E_1 such that $E_i\epsilon M$, $\sum_{j=1}E_j=1$, $E_iE_j=0$ if $i\neq j$, $D_M(E_j)=1/n$. Since $D_{M'}(E_1')=1/n$, $D_{M'}(1)=\alpha\geqq 1$, there exist n projections $E_j'\ \epsilon M'$, $j=1, \cdots, n$, with the first equal to E_1' and such that $E_i'E_j'=0$ if $i\neq j$ and $D_{M'}(E_j')=1/n$. Since $D_M(E_1)=D_M(E_j)$, there is a partially isometric operator $W_j\epsilon M$ with initial set the range of E_1 and final set the range of E_j (cf. R.O., Definition 4.3.1). Similarly we have a $W_j'\ \epsilon M'$ with initial set the range of E_1' and final set E_j'. The relations

$W_j^*W_j=E_1$, $W_jW_j^*=E_j$, $W_j'^*W_j'=E_1'$ and $W_j'W_j'^*=E_j'$ hold, also $W_jE_1=W_j$, $E_jW_j=W_j$, $E_1W_j^*=W_j^*$, $W_j^*E_j=W_j^*$, $W_j'E_1'=W_j'$, $E_j'W_j'=W_j'$, $E_1'W_j'^*=W_j'^*$, $W_j'^*E_j'=W_j'^*$ (cf. R.O., Lemma 4.3.1).

Let $W_j'W_jE_1f=f_j$, $E_1f=f_1$. Then $E_jf_j=E_jW_j'W_jE_1f=W_j'E_jW_jE_1f=W_j'W_jE_1f=f_j$ and $E_j'f_j=E_j'W_j'W_jE_1f=W_j'W_jE_1f=f_j$. Thus $E_jf_j=f_j=E_j'f_j$.

Let $g=\sum_{j=1}^n f_j$, then $E_ig=E_i\sum_{j=1}^n f_j=E_i\sum_{j=1}^n E_jf_j=\sum_{j=1}^n E_iE_jf_j=E_if_i=f_i$. Similarly $E_i'g=f_i$. Now for any $A\epsilon\mathbf{M}$ if either $i\neq j$ or $k\neq l$, then E_iAE_kg is orthogonal to E_jAE_lg. For inasmuch as $E_kg=f_k=E_k'g$, we have,

$$\begin{aligned}(E_iAE_kg,\,E_jAE_lg) &= (E_jE_iAE_kg,\,AE_lg) &&= \delta_j^i(E_iAE_kg,\,AE_lg)\\ &= \delta_j\,(E_iAE_k'g,\,AE_l'g) &&= \delta_j^i(E_k'E_iAg,\,E_l'Ag)\\ &= \delta_j^i\,(E_l'E_k'E_iAg,\,Ag) &&= \delta_j^i\delta_k^l(E_k'E_iAg,\,Ag).\end{aligned}$$

These results imply that if $A\epsilon\mathbf{M}$, then

$$\begin{aligned}\|Ag\|^2 &= \left\|\left(\sum_{i=1}^n E_i\right)A\left(\sum_{j=1}^n E_j\right)g\right\|^2=\left\|\sum_{i=1}^n\sum_{j=1}^n E_iAE_jg\right\|^2=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jg\|^2\\ &=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jE_jg\|^2=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jf_j\|^2=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jW_j'W_jf_1\|^2\\ &=\sum_{i=1}^n\sum_{j=1}^n\|W_j'E_iAE_jW_jf_1\|^2=\sum_{i=1}^n\sum_{j=1}^n\|W_j'E_1'E_iAE_jW_jf_1\|^2\\ &=\sum_{i=1}^n\sum_{j=1}^n\|E_1'E_iAE_jW_jf_1\|^2=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jW_jE_1'f_1\|^2\\ &=\sum_{i=1}^n\sum_{j=1}^n\|E_iAE_jW_jf_1\|^2=\sum_{i=1}^n\sum_{j=1}^n\|W_i^*E_iAE_jW_jf_1\|^2\\ &=\sum_{i=1}^n\sum_{j=1}^n\|W_i^*E_iAE_jW_jE_1f\|^2=\sum_{i=1}^n\sum_{j=1}^n\|W_i^*E_iAE_jW_jf\|^2,\end{aligned}$$

remembering that W_j' is isometric on the range of E_1', W_i^* on the range of E_i and $W_jE_1=W_j$. Substituting A^* for A and interchanging i and j we also have

$$\|A^*g\|^2=\sum_{i=1}^n\sum_{j=1}^n\|W_j^*E_jA^*E_iW_if\|^2=\sum_{i=1}^n\sum_{j=1}^n\|(W_i^*E_iAE_jW_j)^*f\|^2.$$

But recalling the properties of W_j and W_i^* we have

$$E_1W_i^*E_iAE_jW_j=W_i^*E_iAE_jW_j=W_i^*E_iAE_jW_jE_1.$$

Thus Lemma 1.4.3 with $A=W_i^*E_iAE_jW_j$ yields $\|(W_i^*E_iAE_jW_j)^*f\|^2\leq K\|W_i^*E_iAE_jW_jf\|^2$ and with this, we obtain from the above equations for $\|Ag\|^2$ and $\|A^*g\|^2$ that $\|A^*g\|^2\leq K\|Ag\|^2$.

This result may be stated as follows.

LEMMA 1.5.1. *If K is any number >1, there exists a $g\epsilon\mathfrak{H}$, $g\neq 0$ such that for every $A\epsilon M$, $\|A^*g\|^2\leqq K\|Ag\|^2$. Obviously we may assume that $\|g\|=1$.*

1.6. We now have

LEMMA 1.6.1. *Let K and g be as in Lemma 1.5.1. Then if E and F, ϵM, are two projections such that $D_M(E)=D_M(F)$, then $K^{-1}(Fg,g)\leqq(Eg,g)\leqq K(Fg,g)$.*

Inasmuch as $D_M(E)=D_M(F)$, there exists a partially isometric operator W such that $W^*W=F$, $WW^*=E$, $W\epsilon M$ (cf. R.O., Definition 8.2.1, Definition 6.1.1, and Lemma 4.3.1). Now $(Fg,\ g)=(W^*Wg,\ g)=(Wg,\ Wg)=\|Wg\|^2$, $(Eg,\ g)=(WW^*g,\ g)=(W^*g,\ W^*g)=\|W^*g\|^2$. Lemma 1.5.1 with $A=W$ implies $(Eg,\ g)\leqq K(Fg,\ g)$. With $A=W^*$, the same lemma implies $(Fg,\ g)\leqq K(Eg,\ g)$ or $K^{-1}(Fg,\ g)\leqq(Eg,\ g)$. We have now shown our lemma.

Suppose $E\epsilon M$ is such that $D_M(E)=1/m$, where m is an integer. Then there exist $m-1$ projections $E_2,\cdots,E_m$, such that when we let $E_1=E$, we have $\sum_{j=1}^{m}E_j=1$, $E_iE_j=0$, for $i\neq j$, $E_j\epsilon M$, $D_M(E_j)=1/m$. Now returning to the g of Lemmas 1.5.1 and 1.6.1, we have

$$1=\|g\|^2=(g,g)=\left(\sum_{j=1}^{m}E_jg,\ g\right)=\sum_{j=1}^{m}(E_jg,\ g).$$

Let $(E_1g,\ g)=\alpha$. Then Lemma 1.6.1 implies that $K\alpha\geqq(E_jg,\ g)\geqq K^{-1}\alpha$, for every j. Summing over j, gives $Km\alpha\geqq 1\geqq K^{-1}m\alpha$ or $K\alpha\geqq 1/m\geqq K^{-1}\alpha$ which is the same as

$$(+)\qquad K(Eg,\ g)\geqq D_M(E)\geqq K^{-1}(Eg,\ g),$$

since $D_M(E)=1/m$.

Let us study the class Σ of E's in M which satisfy the equation $(+)$. Now if $F_1,\cdots,F_q$ or $F_1,F_2,\cdots$ satisfy $(+)$ and $F_iF_j=0$, $i\neq j$, then $\sum_{j=1}^{q}F_j$ or $\sum_{j=1}^{\infty}F_j$ also satisfies $(+)$. But if $F\epsilon M$ is such that $D_M(F)=q/p$, $F=\sum_{j=1}^{q}F_j$ for mutually orthogonal F_j's ϵM with $D_M(F_j)=1/p$ and hence satisfying $(+)$. Thus if $D_M(F)=q/p$, F satisfies $(+)$.

Let E be any projection of M, $D_M(E)=\alpha$. Let $\{\alpha_i\}$, $\alpha_i\geqq 0$ be a sequence of rational numbers, $\sum_{i=1}^{\infty}\alpha_i=\alpha$. Then there exists a set of mutually orthogonal projections $\{E_i\}$ such that $E_i\leqq E$, $E_i\epsilon M$, $D_M(E_i)=\alpha_i$. To show this, suppose $E_1,\cdots,E_{i-1}$ have been chosen with these properties. Then $D_M(E-\sum_{j=1}^{i-1}E_j)=\sum_{j=i}^{\infty}\alpha_j\geqq\alpha_i$. Hence an E_i can be chosen in such a way that $E_i\epsilon M$ is $\leqq E-\sum_{j=1}^{i-1}E_j$ with $D_M(E_i)=\alpha_i$. Now $D_M(E-\sum_{i=1}^{\infty}E_i)=\alpha-\sum_{i=1}^{\infty}\alpha_i=0$ or $E=\sum_{i=1}^{\infty}E_i$. Since each E_i satisfies $(+)$, E does too. Since E is arbitrary we have

LEMMA 1.6.2. *Let K and g be as in Lemma 1.5.1, then for every $E\epsilon M$*

$$K(Eg,\ g)\geqq D_M(E)\geqq K^{-1}(Eg,\ g).$$

1.7. From Lemma 1.6.2 we can conclude, using Lemma 1.4.2, that if $A \epsilon \boldsymbol{M}$ is positive definite, then

$$K(Ag, g) \geqq Tr_M(A) \geqq K^{-1}(Ag, g).$$

We can state our result as follows.

THEOREM I. *Let* $\boldsymbol{M}$ *be a factor in case* II_1. *Let* $\boldsymbol{M}$ *and* $\boldsymbol{M}'$ *be such that when* D_M *and* $D_{M'}$ *are normalized in such a way that the range of* D_M *is the interval* $(0, 1)$ *and* $C=1$ (cf. §1.1), *then the range of* $D_{M'}$ *is an interval* $(0, \alpha)$ *with* $0<\alpha\leqq\infty$, *then* $\alpha\geqq 1$. *Then to every* $K>1$, *there exists a* $g \epsilon \mathfrak{H}$ *such that for every positive definite* $A \epsilon \boldsymbol{M}$,

$$K(Ag, g) \geqq Tr_M(A) \geqq K^{-1}(Ag, g).$$

CHAPTER II. IMMEDIATE CONSEQUENCES

2.1. We are now able to show that Tr_M has the following two properties when $\boldsymbol{M}$ is in a finite case (cf. R.O., Theorem VIII); that is, when $\boldsymbol{M}$ is in a case I_n $(n=1, 2, \cdots)$ or II_1.

PROPERTY I. *For all Hermitian* A *and* $B \epsilon \boldsymbol{M}$

$$Tr_M(A + B) = Tr_M(A) + Tr_M(B).$$

PROPERTY II. $Tr_M(A)$ *is weakly continuous if* A *is subject to either of these conditions*:

(i) A *is uniformly bounded* (*that is,* $|||A|||\leqq D$ *for some fixed* D);

(ii) A *is definite.*

These properties are obviously independent of the normalization of $D_M(E)$ and $Tr_M(A)$.

If $\boldsymbol{M}$ is in cases I_n, $n=1, 2, \cdots$, these are of course well known results about $Tr_M(A)$ (cf. R.O., p. 220). So we may assume $\boldsymbol{M}$ to be in case II_1.

Then we have in the normalization of R.O., Theorem VIII, three possibilities for the factorization $\boldsymbol{M}, \boldsymbol{M}'$: $\mathrm{II}_1, \mathrm{II}_\infty$; $\mathrm{II}_1, \mathrm{II}_1$ with $C\leqq 1$; $\mathrm{II}_1, \mathrm{II}_1$ with $C\geqq 1$. The two first ones correspond to $\alpha\geqq 1$ in the normalization of §1.1, and since we shall only use Theorem I, from §1.1, we may treat these two cases together.

We therefore consider first the conditions under which Theorem I holds: the normalization of §1.1, and $\alpha\geqq 1$. We obtain an equivalent statement of Theorem I with $A \epsilon \boldsymbol{M}$, Hermitian and not necessarily definite.

Let K, g be as in Theorem I, then $K^{-1}(g, g)=K^{-1}\|g\|^2\leqq Tr_M(1)=1$, or $\|g\|^2\leqq K$. Furthermore if $A \epsilon \boldsymbol{M}$ is Hermitian, then $A=A_1-A_2$, where A_1 and A_2 are both positive definite and $A_1A_2=0=A_2A_1$. (Take $A_1=\frac{1}{2}(A+|A|)$,

$A_2=\frac{1}{2}(|A|-A)$ (cf. (15), Chapter VI, or (19), p. 203 ff.).) Since A_1 and A_2 commute, $Tr_M(A)=Tr_M(A_1)-Tr_M(A_2)$ by R.O., Lemma 15.3.4. It is a consequence of Theorem 1 that, if $\epsilon=K-1$,

$$|\,Tr_M(A_1) - (A_1g, g)\,| \leqq \epsilon(A_1g, g); \quad |\,Tr_M(A_2) - (A_2g, g)\,| \leqq \epsilon(A_2g, g).$$

These with the above equation for $Tr_M(A)$ imply

$$|\,Tr_M(A) - (Ag, g)\,| \leqq \epsilon((A_1g, g) + (A_2g, g)).$$

But the bounds of A_1 and A_2 are each not more than that of A. Also $\|g\|^2\leqq 1+\epsilon$. So if A has the bound D, we have

$$|\,Tr_M(A) - (Ag, g)\,| \leqq 2\epsilon(1 + \epsilon)D.$$

Let A, B, $A+B$ have the respective bounds D_1, D_2, D_3. Substituting in (1) each of these operators, inasmuch as $((A+B)g, g)=(Ag, g)+(Bg, g)$, we get

$$|\,Tr_M(A + B) - Tr_M(A) - Tr_M(B)\,| \leqq 2\epsilon(1 + \epsilon)(D_1 + D_2 + D_3),$$

which since ϵ may be taken arbitrarily small, implies Property I.

We turn our attention to Property II, (i). Consider the set Σ of all Hermitian $A\epsilon M$, with a bound less than or equal to a fixed D. We will show that to every $A\epsilon\Sigma$ and to every $\eta>0$, there exists a weak neighborhood of A, $U(A; g; g; \eta/3)$ such that for all $B\epsilon\Sigma$ and $B\epsilon U(A; g; g; \eta/3)$ (i.e., $|((B-A)g, g)|<\eta/3$) we have $|Tr_M(B)-Tr_M(A)|<\eta$. Letting ϵ be such that $\epsilon(1+\epsilon)D\leqq\eta/3$, as above we can find a g such that for A and every $B\epsilon\Sigma$, $|Tr_M(A)-(Ag, g)|\leqq\epsilon(1+\epsilon)D\leqq\eta/3$, $|Tr_M(B)-(Bg, g)|\leqq\eta/3$. In particular for $B\epsilon U(A; g; g; \eta/3)$ (cf. above) which means that we also have $|((B-A)g, g)|<\eta/3$, these inequalities imply that $|Tr_M(A)-Tr_M(B)|<\eta$.

Thus we have shown the weak neighborhood continuity of $Tr_M(A)$ with of course a restriction. From this the sequential continuity follows immediately. (If $A_n\to A$, then the A_n are uniformly bounded and for any $\eta>0$, almost all the A_n must be in the above given neighborhood.) In this situation the two kinds of continuity are equivalent (cf. (18), pp. 383–384), but we will use only the sequential.

$Tr_M(A)$ is also weakly continuous when considered only for the positive definite $A\epsilon M$. For suppose A is such and an $\epsilon>0$ is given. Let the K of Theorem I be chosen in such a way that $1<K\leqq 2$, and $(K-1)Tr_M(A)<\epsilon/5$, and let g correspond to this K as there. Let η be chosen in such a way that $\eta\leqq\epsilon/5$. Then if a positive definite $B\epsilon M$ is in $U(A; g; g; \eta)$, we have $|Tr_M(A)-Tr_M(B)|<\epsilon$ (because $|((B-A)g, g)|<\eta\leqq\epsilon/5$, $|Tr_M(A)-(Ag, g)|$

$\leqq(K-1)Tr_M(A)<\epsilon/5$, and, in addition, $|Tr_M(B)-(Bg,g)|\leqq(K-1)(Bg,g)$ $\leqq(K-1)((Ag,g)+\eta)\leqq(K-1)(Tr_M(A)+\epsilon/5+\eta)\leqq 3\epsilon/5)$.

Thus we have settled the factorizations II_1, II_∞ and II_1, II_1 with $C\leqq 1$. But we still must consider the case II_1, II_1 with $C>1$. Let m be an integer such that $C/m\leqq 1$, E_m an m-dimensional unitary space. Form $E_m\otimes\mathfrak{H}$ (cf. R.O., Theorem I). Let $\boldsymbol{N}$ be the ring of operators on E_m, $\boldsymbol{N}'$ the corresponding set of operators on $E\otimes\mathfrak{H}$ (cf. R.O., Lemma 2.3.1 and Lemma 2.3.2), $\boldsymbol{B}$ the set of all operators in $\mathfrak{H}$, $\boldsymbol{B}^{(2)}$ the corresponding set in $E\otimes\mathfrak{H}$.

Under the above correspondence of $\boldsymbol{B}$ and $\boldsymbol{B}^{(2)}$, $\boldsymbol{M}\sim\boldsymbol{M}^{(2)}$, $\boldsymbol{M}'\sim\boldsymbol{M}'^{(2)}$. The first and hence each of the others is a full ring isomorphism (R.O., Lemma 2.3.4 and (22), §4). Thus $\boldsymbol{M}^{(2)}$ and $\boldsymbol{M}'^{(2)}$ are in case II_1 (R.O., Theorem IX). Now let $\phi_1,\cdots,\phi_m$ be a complete orthonormal set in E_m and let us think of the set up of R.O., §2.4. By R.O., Lemma 2.4.5, we have $\boldsymbol{M}^{(2)\prime}=R(\boldsymbol{N}^{(1)},\boldsymbol{M}'^{(2)})$ and by R.O., Lemma 11.5.2, $R(\boldsymbol{N}^{(1)},\boldsymbol{M}'^{(2)})$ is in case II_1 since as we have mentioned before, $\boldsymbol{M}'^{(2)}$ is in this case.

Thus the coupled factorization, $\boldsymbol{M}^{(2)}$, $\boldsymbol{M}^{(2)\prime}$, is in case II_1, II_1 and $\boldsymbol{M}^{(2)}$ is fully ring isomorphic to $\boldsymbol{M}$. Furthermore such an isomorphism preserves D_M in the standard normalization (R.O., Theorem IX) and hence the trace. Now suppose we have shown that for $\boldsymbol{M}^{(2)}$, $\boldsymbol{M}^{(2)\prime}$ we have $C^{(2)}\leqq 1$. Then $\boldsymbol{M}^{(2)}$ has Properties I and II, by the above and $\boldsymbol{M}$ must have them to be the isomorphism. (As m is finite, even weak operator topology is conserved under the mapping $\boldsymbol{B}\sim\boldsymbol{B}^{(2)}$.)

Therefore it remains to show that $C^{(2)}\leqq 1$. Consider the closed linear manifold, $(\phi_1\otimes f; f\epsilon\mathfrak{H})=$ (say) $\mathfrak{M}$, in $E_m\otimes\mathfrak{H}$. The correspondence of $\mathfrak{M}$ and $\mathfrak{H}$ given by $\phi_1\otimes f\sim f$ is unitary. Under it, $\boldsymbol{M}^{(2)}_{\mathfrak{M}}$ and $R(\boldsymbol{N}^{(1)},\boldsymbol{M}'^{(2)})_{\mathfrak{M}}$ (cf. R.O., Definition 11.3.1) are unitarily equivalent to $\boldsymbol{M}$ and $\boldsymbol{M}'$ (cf. R.O., §2.4). Thus the C for $\boldsymbol{M}$, $\boldsymbol{M}'$ is the same as that of $\boldsymbol{M}^{(2)}_{\mathfrak{M}}$ and $R(\boldsymbol{N}^{(1)},\boldsymbol{M}'^{(2)})_{\mathfrak{M}}$. Then from the proof of Lemma 11.4.3, we can conclude that $C=(D_{M^{(2)}}(\mathfrak{H})/D_{M^{(2)\prime}}(\mathfrak{M}))C^{(2)}=mC^{(2)}$ or $C^{(2)}=C/m\leqq 1$.

2.2. Previously $Tr_M(A)$ has only been considered for Hermitian $A\epsilon\boldsymbol{M}$. We now define it for all $A\epsilon\boldsymbol{M}$.

Definition 2.2.1. *If $A\epsilon\boldsymbol{M}$, then $A=B+iC$, where $B=\frac{1}{2}(A+A^*)$, $C=-\frac{1}{2}i(A-A^*)$ are Hermitian, and we define $Tr_M(A)=Tr_M(B)+iTr_M(C)$.*

$Tr_M(A)$ has the following property.

Property III. (*) *$Tr_M(A)$ agrees with the previous definition of $Tr_M(A)$ if A is Hermitian.*

(**) *For uniformly bounded or for definite A's, $Tr_M(A)$ is weakly continuous.*

(i) $Tr_M(1)=1$;
(ii) $Tr_M(\rho A)=\rho Tr_M(A)$, *$\rho$ complex*;
(iii) $Tr_M(A+B)=Tr_M(A)+Tr_M(B)$;
(iv) $Tr_M(A)\geqq 0$, *if A is definite*;
(iv)′ *If A is definite and $\neq 0$, $Tr_M(A)>0$*;
(v) $Tr_M(A^*)=\overline{Tr_M(A)}$;
(vi) $Tr_M(AB)=Tr_M(BA)$;
(vi)′ $Tr_M(U^{-1}AU)=Tr_M(A)$, *if U is unitary.*

Now, $(*)$ and (v) are immediate consequences of the definition. (i), (iv), and (iv)′ follow from $(*)$ and R.O., Theorem XIII. A^* is a weakly continuous additive function of A. Thus B and C are too, and we may infer (iii) and $(**)$ from Properties I and II respectively. To show (ii) we first note that if A is Hermitian by Definition 2.2.1, $Tr_M(iA)=iTr_M(A)$ and if a is real by $(*)$ $Tr_M(aA)=aTr_M(A)$. Then (ii) can be shown by a calculation involving (iii).

To show (vi) suppose A and $B\epsilon M$ are Hermitian. Let $C=A+iB$, $C^*=A-iB$. Then $Tr_M(CC^*)=Tr_M(C^*C)$ (by $(*)$, cf. proof of Lemma 1.4.3, equation (β)). Substituting for C, expanding the products according to the distributive law and collecting terms, we get $2i(Tr_M(AB)-Tr_M(BA))=0$ or $Tr_M(AB)=Tr_M(BA)$. From this we can show (vi) in general by using Definition 2.2.1, (ii) and (iii). (vi)′ follows from (vi) by letting $A=U^{-1}A$, $B=U$ and using the associative law.

Property IV. (i)–(iv), (v), (vi)′ *of Property* III *characterize $Tr_M(A)$ uniquely.*

Let $Tr'(A)$ have the listed properties. Then if A is Hermitian, $Tr'(A)$ is real by (v). This and (i)–(iv), (vi)′ imply that for $A\epsilon M$ and Hermitian $Tr'(A)=Tr_M(A)$ (R.O., Theorem XIII). By (ii) and (iii), we then have that for an arbitrary $C\epsilon M$, $C=A+iB$, A and B Hermitian, $Tr'(C)=Tr'(A+iB)$ $=Tr'(A)+Tr'(iB)=Tr'(A)+iTr'(B)=Tr_M(A)+iTr_M(B)=Tr_M(A+iB)$ $=Tr_M(C)$ or $Tr'(C)=Tr_M(C)$.

Chapter III. The exact form of $Tr_M(A)$ (for $\alpha\geqq 1$)

3.1. We again assume M in case II_1. Some lemmas about families of projections and spectral forms in M are needed. Of these the two last ones have some interest of their own.

Lemma 3.1.1. *For any two projections E, $F\epsilon M$ with $E\leqq F$, and any $\alpha\geqq D_M(E)$, $\leqq D_M(F)$, a projection $G\epsilon M$ with $E\leqq G\leqq F$, $D_M(G)=\alpha$ exists.*

$F-E$ is a projection, and $0\leqq\alpha-D_M(E)\leqq D_M(F)-D_M(E)=D_M(F-E)$

$\leqq 1$. Thus a projection $G' \epsilon M$ with $D_M(G') = \alpha - D_M(E)$ exists (because M is in case II_1). So $D_M(G') \leqq D_M(F-E)$ and therefore a projection $G'' \epsilon M$ with $G'' \leqq F-E$, $D_M(G'') = D_M(G') = \alpha - D_M(E)$ exists. G'' is thus orthogonal to E, and therefore $G''+E$ is a projection and $D_M(G''+E) = D_M(G'') + D_M(E) = \alpha$. Besides, $G''+E \geqq E$ and $G''+E \leqq (F-E)+E = F$. So $G = G''+E$ has the desired properties.

LEMMA 3.1.2. *For any two projections E, $F \epsilon M$ with $E \leqq F$ a family of projections $G(\alpha) \epsilon M$ defined for all α with $D_M(E) \leqq \alpha \leqq D_M(F)$ exists, which possesses the following properties:*

(i) *$G(\alpha) = E$ or F if $\alpha = D_M(E)$ or $D_M(F)$ respectively.*

(ii) *$\alpha \leqq \beta$ implies $G(\alpha) \leqq G(\beta)$.*

(iii) *$D_M(G(\alpha)) = \alpha$.*

Choose a sequence $\rho_1, \rho_2, \cdots$ which lies and is dense in the interval $D_M(E) \leqq \alpha \leqq D_M(F)$, with $\rho_1 = D_M(E)$, $\rho_2 = D_M(F)$.

We will now define a sequence of projections $G(\rho_1), G(\rho_2), \cdots, \epsilon M$, so that (i)–(iii) hold for $\alpha = \rho_1, \rho_2, \cdots$. Put first $G(\rho_1) = E$, $G(\rho_2) = F$; then (i)–(iii) hold for $\alpha = \rho_1, \rho_2$. Assume now that for a $j = 3, 4, \cdots$ the $G(\rho_1), \cdots, G(\rho_{j-1})$ are already defined, so that (i)–(iii) holds for $\alpha = \rho_1, \cdots, \rho_{j-1}$, we will now define $G(\rho_j)$ without violating (i)–(iii).

Consider the ρ_i, $i = 1, \cdots, j-1$, with $\rho_i \leqq \rho_j$ (they exist: $\rho_1 = D_M(E) \leqq \rho_j$); let $\rho_{i'}$ be the greatest one. Consider the ρ_i, $i = 1, \cdots, j-1$, with $\rho_i \geqq \rho_j$ (they exist: $\rho_2 = D_M(F) \geqq \rho_j$); let $\rho_{i''}$ be the smallest one. Thus $\rho_{i'} \leqq \rho_{i''}$ and so $G(\rho_{i'}) \leqq G(\rho_{i''})$. Besides $D_M(G(\rho_{i'})) = \rho_{i'} \leqq \rho_j \leqq \rho_{i''} = D_M(G(\rho_{i''}))$. So Lemma 3.1.1 can be applied to $E = G(\rho_{i'})$, $F = G(\rho_{i''})$, $\alpha = \rho_j$, and we define $G(\rho_j) = G$.

Thus $G(\rho_{i'}) \leqq G(\rho_j) \leqq G(\rho_{i''})$, therefore $\rho_i \leqq \rho_j$, $i = 1, \cdots, j-1$, implies $\rho_i \leqq \rho_{i'}$, $G(\rho_i) \leqq G(\rho_{i'}) \leqq G(\rho_j)$ and $\rho_i \geqq \rho_j$, $i = 1, \cdots, j-1$, implies $\rho_i \geqq \rho_{i''}$, $G(\rho_i) \geqq G(\rho_{i''}) \geqq G(\rho_j)$. Besides $D_M(G(\rho_j)) = \rho_j$. So (i)–(iii) hold for $\alpha = \rho_1, \cdots, \rho_{j-1}, \rho_j$, too.

Thus all $G(\rho_1), G(\rho_2), \cdots$ are defined. Owing to (ii), $\lim_{\rho_i \to \alpha, \rho_i \geqq \alpha} G(\rho_i)$ exists for all α with $D_M(E) = \rho_1 \leqq \alpha \leqq \rho_2 = D_M(F)$. If α is equal to a ρ_j, then this limit is meant to denote the value at ρ_j. So we can extend the definition of $G(\alpha)$ to all above α by defining

$$G(\alpha) = \lim_{\rho_i \to \alpha,\, \rho_i \geqq \alpha} G(\rho_i).$$

Now (i) remains true, and (ii), (iii) extend by continuity to all above α.

LEMMA 3.1.3. *Let $A \epsilon M$ be a definite operator with bound $\leqq 1$. Then there exists a monotonous non-decreasing and right semi-continuous function $\phi(\alpha)$ for $0 \leqq \alpha \leqq 1$ with $0 \leqq \phi(\alpha) \leqq 1$ and a resolution of identity $E(\alpha) \epsilon M$* (cf. (16), p. 92, or R.O., Definition 15.1.1) *with the following properties:*

(i) $E(0)=0$, $E(1)=1$,

(ii) $D_M(E(\alpha))=\alpha$,

(iii) $(Af, g)=\int_0^1\phi(\alpha)d(E(\alpha)f, g)$ *for any* f, g, *or symbolically*,

$$A = \int_0^1 \phi(\alpha)dE(\alpha).$$

Let $F(\lambda)$ be the resolution of identity corresponding to A. As A is definite, so $F(0)=0$; as the bound of A is $\leqq 1$, so $F(1)=1$.

Consider now the two following functions:

$$\psi(\lambda) = D_M(F(\lambda)),$$
$$\phi(\alpha) = \underset{0\leqq\lambda\leqq 1,\ \psi(\lambda)\leqq\alpha}{\text{l.u.b.}} \quad \lambda.$$

$\psi(\lambda)$ is monotonous non-decreasing and right semi-continuous in $0\leqq\lambda\leqq 1$, because $F(\lambda)$ is a resolution of identity, and besides $\psi(0)=0$, $\psi(1)=1$. Thus $\phi(\alpha)$ is monotonous non-decreasing and right semi-continuous in $0\leqq\alpha\leqq 1$, and its values are all in $0\leqq\lambda\leqq 1$. One verifies these facts, as well as those which follow, easily; besides, their discussion may be found in (19) (numbers in parentheses refer to the bibliography in R.O., pp. 125–126) on p. 193. (Our $\psi(\lambda)$, $\phi(\alpha)$ corresponds to the $\phi(a)$, $\psi(b)$ there, thus they are both "m-functions," and $\phi(\alpha)$ is the "measure function" of $\psi(\lambda)$.) The relation between $\phi(\alpha)$ and $\psi(\lambda)$ is symmetric (cf. loc. cit., $\psi(\lambda)$ is the "measure function" of $\phi(\alpha)$, we must define for the empty set in $0\leqq\lambda\leqq 1$ the l.u.b. 0); that is,

$$\psi(\lambda) = \underset{0\leqq\alpha\leqq 1}{\text{l.u.b.}} (\phi(\alpha) \leqq \lambda).$$

Finally (cf. loc. cit.)

$$(*) \qquad \psi(\phi(\psi(\lambda))) = \psi(\lambda).$$

That is, $D_M(F(\phi(\psi(\lambda)))=D_M(F(\lambda))$. But $F(\phi(\psi(\lambda)))\gtreqless F(\lambda)$, and therefore this implies

$$(**) \qquad F(\phi(\psi(\lambda))) = F(\lambda).$$

If $\phi(\alpha)<1$, then choose μ with $\phi(\alpha)<\mu\leqq 1$, and let $\mu\rightarrow\phi(\alpha)$. Under these conditions the definition of $\phi(\alpha)$ implies $D_M(F(\mu))>\alpha$, so $D_M(F(\phi(\alpha)))\geqq\alpha$. For $\phi(\alpha)=1$ this holds, too: $D_M(F(\phi(\alpha)))=D_M(F(1))=D_M(1)=1\geqq\alpha$. On the other hand, if $\phi(\alpha)>0$, then a sequence of μ with $0\leqq\mu\leqq\phi(\alpha)$, $\mu\rightarrow\phi(\alpha)$ and $D_M(F(\mu))\leqq\alpha$ exists, so $D_M(F(\phi(\alpha)-0))\leqq\alpha$. This holds for $\phi(0)=0$, too, if we identify $F(0-0)$ with $F(0)=0$: $D_M(F(0-0))=D_M(0)=0\leqq\alpha$. So we have

$$(\#) \qquad D_M(F(\phi(\alpha) - 0)) \leqq \alpha \leqq D_M(F(\phi(\alpha))).$$

Form now for every $0 \leqq \lambda \leqq 1$, $E = F(\lambda - 0)$, $F = F(\lambda)$ and apply Lemma 3.1.2. A family of projections $G(\alpha) \epsilon \boldsymbol{M}$, $D_M(F(\lambda-0)) \leqq \alpha \leqq D_M(F(\lambda))$, results. Choose for every $0 \leqq \lambda \leqq 1$ such a family, and hold it fixed: $G(\lambda, \alpha)$. (If $F(\lambda-0) = F(\lambda)$, then necessarily $\alpha = D_M(F(\lambda))$, and $G(\lambda, \alpha) = F(\lambda)$. So only the λ with $F(\lambda-0) \neq F(\lambda)$, the point-proper values of A, are important. Their set is finite or enumerably infinite.)

Now define for every $0 \leqq \alpha \leqq 1$

$$(\S) \qquad E(\alpha) = G(\phi(\alpha), \alpha)$$

(this expression has a meaning owing to (#)).

We have now by definition $D_M(E(\alpha)) = \alpha$, so condition (ii) is satisfied. For $\alpha = 0, 1$, this gives $E(0) = 0$, $E(1) = 1$, so condition (i) is satisfied, too. If $\alpha \leqq \beta$, then $\phi(\alpha) \leqq \phi(\beta)$. The relation $\phi(\alpha) = \phi(\beta)$ gives

$$E(\alpha) = G(\phi(\alpha), \alpha) = G(\phi(\beta), \alpha) \leqq G(\phi(\beta), \beta) = E(\beta),$$

while $\phi(\alpha) < \phi(\beta)$ gives

$$E(\alpha) = G(\phi(\alpha), \alpha) \leqq F(\phi(\alpha)) \leqq F(\phi(\beta) - 0) \leqq G(\phi(\beta), \beta) = E(\beta).$$

So we see that

$$(\cdot) \qquad \alpha \leqq \beta \quad \text{implies} \quad E(\alpha) \leqq E(\beta).$$

Furthermore $\beta > \alpha$ gives

$$D_M(E(\beta) - E(\alpha)) = D_M(E(\beta)) - D_M(E(\alpha)) = \beta - \alpha,$$

so $\lim_{\beta \to \alpha, \beta > \alpha} D_M(E(\beta) - E(\alpha)) = 0$. But

$$\lim_{\beta \to \alpha, \beta > \alpha} D_M(E(\beta) - E(\alpha)) = D_M\left(\lim_{\beta \to \alpha, \beta > \alpha} E(\beta) - E(\alpha)\right) = D_M(E(\alpha + 0) - E(\alpha))$$

so $D_M(E(\alpha+0) - E(\alpha)) = 0$, $E(\alpha+0) = E(\alpha)$, or

$$(\cdot\,\cdot) \qquad \lim_{\beta \to \alpha, \beta > \alpha} E(\beta) = E(\alpha).$$

(i), $(\cdot)$, $(\cdot\,\cdot)$ state that $E(\alpha)$ is a resolution of unity.

It remains for us to prove (iii).

Consider the expression $\int_0^1 \phi(\alpha) d(\|E(\alpha)f\|^2)$. $\phi(\alpha)$ and $\|E(\alpha)f\|^2$ are both monotonous non-decreasing functions of α if α increases from 0 to 1, then these functions increase from $\phi(0)$ to $\phi(1) = 1$ and from 0 to $\|f\|^2$ respectively. Thus our integral is, by partial integration, equal to $\|f\|^2 - \int_0^1 \|E(\alpha)f\|^2 d\phi(\alpha)$. We may now introduce a new variable of integration: $\lambda = \phi(\alpha)$. Then we have (cf. loc. cit., p. 198) $\int_0^1 \|E(\alpha)f\|^2 d\phi(\alpha) = \int_0^1 \|E(\psi(\lambda))f\|^2 d\lambda$. But by (§) $E(\psi(\lambda)) = G(\phi(\psi(\lambda)), \psi(\lambda))$, and by (*) $\psi(\phi(\psi(\lambda))) = D_M(F(\phi(\psi(\lambda))))$. So the

definition of $G(\mu, \alpha)$ gives $G(\phi(\psi(\lambda)), \psi(\lambda)) = F(\phi(\psi(\lambda)))$. By $(**)$ this is $F(\lambda)$, so we have $E(\psi(\lambda)) = F(\lambda)$. Thus our original expression is equal to $\|f\|^2 - \int_0^1 \|F(\lambda)f\|^2 d\lambda$, and this becomes by another partial integration $\int_0^1 \lambda d\|F(\lambda)f\|^2$, that is, (Af, f). Thus we have proved:

$$(Af, f) = \int_0^1 \phi(\alpha) d\|E(\alpha)f\|^2.$$

Replacing herein f by $f \pm g/2$, and subtracting, gives the real part of (iii). Now replacing f, g by if, g gives the imaginary part of (iii). Thus the proof is completed.

Finally we need the following lemma.

LEMMA 3.1.4. *Let $E(\alpha) \epsilon \boldsymbol{M}$ be a resolution of unity in $a \leqq \alpha \leqq b$, $E(a) = 0$, $E(b) = 1$. Then a bounded Hermitian operator $B \epsilon \boldsymbol{M}$ with $(Bf, g) = \int_a^b \lambda d(E(\lambda)f, g)$, symbolically $B = \int_a^b \lambda dE(\lambda)$, exists. For every bounded Baire function $\psi(\alpha)$ an operator $C = \psi(B) \epsilon \boldsymbol{M}$ with $(Cf, g) = \int_a^b \psi(\lambda) d(E(\lambda)f, g)$ symbolically $C = \int_a^b \psi(\lambda) dE(\lambda)$ exists. If $0 \leqq \psi(\lambda) \leqq 1$, then C is Hermitian, definite, and of bound 1. For all $\psi(\lambda)$*

$$Tr_{\boldsymbol{M}}(C) = \int_a^b \psi(\lambda) dD_{\boldsymbol{M}}(E(\lambda)).$$

The existence of the bounded B is a well known fact about spectral forms, as all $E(\lambda) \epsilon \boldsymbol{M}$ so $B \epsilon \boldsymbol{M}$ (see (18), p. 389, Theorem 1). $C = \psi(B)$ exists and is bounded ((19), p. 205), and $B \epsilon \boldsymbol{M}$ implies $C \epsilon \boldsymbol{M}$ ((19), p. 213, Theorem 6). $0 \leqq \psi(\alpha) \leqq 1$ implies the Hermitian, definite character of C and $1 - C$ ((19), p. 205) (C Hermitian: by Property b; C and $1 - C$ definite: by Property e), with $F(x) = \psi(x)$ or $1 - \psi(x)$ and $G(x) = H(x) = (F(x))^{1/2}$. Thus the bound of C is $\leqq 1$. (Alternatively (16), p. 113, Theorem 4*, could be used.)

It remains to prove the final formula for the trace. If it holds for $\psi_1(\lambda), \psi_2(\lambda), \cdots$, and the $\psi_n(\lambda)$, $n = 1, 2, \cdots$ are uniformly bounded and everywhere convergent in $a \leqq \lambda \leqq b$, then it holds for $\psi(\lambda) = \lim_{n\to\infty} \psi_n(\lambda)$ too: $\psi(B) = \text{strong} \lim_{n\to\infty} \psi_n(B)$, the $\psi_n(B)$ being uniformly bounded, by (19), p. 205, Property h), and $Tr_{\boldsymbol{M}}$ is continuous for strong (even for weak) convergence by Property III, $(**)$. Thus our relation holds for all bounded Baire functions $\psi(\lambda)$ if it holds for all continuous functions $\psi(\lambda)$. And it holds for these, if it holds for all intervalwise constant functions $\psi(\lambda)$. But these are linear aggregates of functions $\psi_{c,d}(\lambda) = 1$ for $c < \lambda \leqq d = 0$ otherwise, where $a \leqq c \leqq d \leqq b$. So it suffices to consider these. Now clearly

$$\psi_{c,d}(B) = E(d) - E(c),$$
$$Tr_{\boldsymbol{M}}(\psi_{c,d}(B)) = D_{\boldsymbol{M}}(E(d) - E(c)) = D_{\boldsymbol{M}}(E(d)) - D_{\boldsymbol{M}}(E(c)),$$

and

$$\int_a^b \psi_{c,d}(\alpha)dD_M(E(\alpha)) = D_M(E(d)) - D_M(E(c)),$$

completing the proof.

3.2. We return now to the normalization of §1.1, and assume that an $\boldsymbol{M}$ in case II_1 is given with $\alpha \geqq 1$.

Consider a g and K which are related as in Theorem I. Our objective is to prove Theorem II below, but we begin by considering these two maximum problems.

(A) For a given λ with $0 \leqq \lambda \leqq 1$ consider all projections $F \epsilon \boldsymbol{M}$ with $D_M(F) = \lambda$, and the corresponding values of (Fg, g). Prove that these (Fg, g) possess a maximum m_a which they assume for a certain $F = F_0$ from the above class.

(B) For a given λ with $0 \leqq \lambda \leqq 1$ consider all definite operators $B \epsilon \boldsymbol{M}$ of bound $\leqq 1$ with $Tr_M(B) = \lambda$, and the corresponding values of (Bg, g). Prove that these (Bg, g) possess a maximum m_b which they assume for a certain $B = B_0$ from the above class.

LEMMA 3.2.1. *Problem* (B) *possesses a solution $B = B_0$, the maximum m_b fulfills $\lambda K^{-1} \leqq m_b \leqq \lambda K$.*

For the B of the class described in Problem (B) we have by Theorem I, $\lambda K^{-1} \leqq (Bg, g) \leqq \lambda K$. So these (Bg, g) possess a l.u.b. m_b' and $\lambda K^{-1} \leqq m_b' \leqq \lambda K$.

Select now a sequence $B_1, B_2, \cdots$ from this class, so that $\lim_{n\to\infty}(B_n g, g) = m_b'$. As the $B_1, B_2, \cdots$ are uniformly bounded, there exists a subsequence $B_{n_1}, B_{n_2}, \cdots$ $(n_1 < n_2 < \cdots)$ such that $B_0 = \text{weak lim } B_{n_i}$ exists. As all B_{n_i} are $\epsilon \boldsymbol{M}$, definite, and of bound $\leqq 1$, the same is true for B_0. As all $Tr_M(B_{n_i}) = \lambda$, so Property II or III, $(**)$ gives $Tr_M(B_0) = \lambda$. Finally $(B_0 g, g) = \lim_{i\to\infty}(B_{n_i} g, g) = \lim_{n\to\infty}(B_n g, g) = m_b'$.

Thus B_0 belongs to our class, and $(B_0 g, g) = m_b'$. Therefore the l.u.b. m_b' is a maximum: $m_b' = m_b$, and $\lambda K^{-1} \leqq m_b \leqq \lambda K$. So the proof is completed.

LEMMA 3.2.2. *The B_0 of Lemma* 3.2.1 *can be chosen as a projection.*

Consider the B_0 of Lemma 3.2.1. By Lemmas 3.1.3 and 3.1.4 we have $B_0 = \phi(B)$, where $B = \int_0^1 \alpha dE(\alpha)$ (symbolically), $E(\alpha)$ being a resolution of unity with $E(\alpha) \epsilon \boldsymbol{M}$, $D_M(E(\alpha)) = \alpha$, and $\phi(\alpha)$ a monotonous non-decreasing and right semi-continuous function.

Consider any Baire function $\psi(\alpha)$, $0 \leqq \alpha \leqq 1$, with $0 \leqq \psi(\alpha) \leqq 1$. Form $\psi(B) = \int_0^1 \psi(\alpha) dE(\alpha)$ (symbolically), using Lemma 3.1.4. Then $\psi(B)$ is $\epsilon \boldsymbol{M}$, definite, and of bound $\leqq 1$, and

$$Tr_M(\psi(B)) = \int_0^1 \psi(\alpha)dD_M(E(\alpha)) = \int_0^1 \psi(\alpha)d\alpha.$$

Thus $\psi(B)$ belongs to the class of Problem (B) if and only if $\int_0^1\psi(\alpha)d\alpha=\lambda$. Besides $(\psi(B)g, g)=\int_0^1\psi(\alpha)d(\|E(\alpha)g\|^2)$.

Apply this to $\psi(\alpha)=\phi(\alpha)$, $\psi(B)=\phi(B)=B_0$. We get $\int_0^1\phi(\alpha)d\alpha=\lambda$, $\int_0^1\phi(\alpha)d(\|E(\alpha)g\|^2)=m_b$. And the maximum property of B_0 gives the following. For every Baire function $\psi(\alpha)$, $0\leqq\alpha\leqq 1$, with $0\leqq\psi(\alpha)\leqq 1$ the equation $\int_0^1\psi(\alpha)d\alpha=\lambda$ implies $\int_0^1\psi(\alpha)d\|E(\alpha)g\|^2\leqq m_b$.

If we had $\phi(1-\lambda)=0$, then the monotony of $\phi(\alpha)$ would give $\phi(\alpha)=0$ for $0\leqq\alpha\leqq 1-\lambda$, and the right semi-continuity $\phi(\alpha)\leqq\frac{1}{2}$ for $1-\lambda<\alpha\leqq 1-\lambda+\epsilon'$ for a suitable $\epsilon'>0$. Besides $\phi(\alpha)$ is always $\leqq 1$. Thus

$$\int_0^1 \phi(\alpha)d\alpha \leqq 0(1-\lambda) + \tfrac{1}{2}\epsilon' + (\lambda-\epsilon')1 = \lambda - \tfrac{1}{2}\epsilon' < \lambda$$

contradicting $\int_0^1\phi(\alpha)d\alpha=\lambda$. Therefore $\phi(1-\lambda)>0$.

If we had $\phi(1-\lambda-0)=1$, then the monotony of $\phi(\alpha)$ would give $\phi(\alpha)=1$ for $1-\lambda\leqq\alpha\leqq 1$ and the definition of $\phi(1-\lambda-0)$, $\phi(\alpha)\geqq\frac{1}{2}$ for $1-\lambda-\epsilon'\leqq\alpha<1-\lambda$, for a suitable $\epsilon'>0$. Besides $\phi(\alpha)$ is always $\geqq 0$. Thus

$$\int_0^1 \phi(\alpha)d\alpha \geqq (1-\lambda-\epsilon')\cdot 0 + \frac{\epsilon'}{2} + \lambda\cdot 1 = \lambda + \frac{\epsilon'}{2} > \lambda$$

contradicting $\int_0^1\phi(\alpha)d\alpha=\lambda$. Therefore $\phi(1-\lambda-0)<1$.

So we can choose a δ, $0<\delta<1$, with $\phi(1-\lambda)>\delta$, $\phi(1-\lambda-0)<1-\delta$. Thus $1-\lambda\leqq\alpha\leqq 1$ implies $(\phi(\alpha)-\delta)/(1-\delta)\geqq(\phi(1-\lambda)-\delta)/(1-\delta)\geqq 0$ and $\leqq(1-\delta)/(1-\delta)=1$, and $0\leqq\alpha\leqq 1-\lambda$ implies $\phi(\alpha)/(1-\delta)\geqq 0$ and $\leqq(1-\delta)/(1-\delta)=1$. Now put $\phi_1(\alpha)=1$ for $1-\lambda\leqq\alpha\leqq 1$ and $\phi_1(\alpha)=0$ for $0\leqq\alpha\leqq 1-\lambda$. Then we have $0\leqq\phi_1(\alpha)\leqq 1$ for all $0\leqq\alpha\leqq 1$; and also, for $\phi_2(\alpha)=(\phi(\alpha)-\delta\phi_1(\alpha))/(1-\delta)$, $0\leqq\phi_2(\alpha)\leqq 1$ for all $0\leqq\alpha\leqq 1$ (we proved this above separately for $1-\lambda\leqq\alpha\leqq 1$ and for $0\leqq\alpha\leqq 1-\lambda$). Besides

$$(*) \qquad \phi(\alpha) = \delta\phi_1(\alpha) + (1-\delta)\phi_2(\alpha).$$

Now clearly $\int_0^1\phi_1(\alpha)d\alpha=\lambda$. This and $\int_0^1\phi(\alpha)d\alpha=\lambda$ and $(*)$ give $\int_0^1\phi_2(\alpha)d\alpha=\lambda$. Therefore $\int_0^1\phi_1(\alpha)d\|E(\alpha)g\|^2\leqq m_b$, $\int_0^1\phi_2(\alpha)d\|E(\alpha)g\|^2\leqq m_b$. But $(*)$ gives

$$\delta\int_0^1\phi_1(\alpha)d\|E(\alpha)g\|^2 + (1-\delta)\int_0^1\phi_2(\alpha)d\|E(\alpha)g\|^2 = \int_0^1\phi(\alpha)d\|E(\alpha)g\|^2 = m_b.$$

Therefore necessarily $\int_0^1\phi_1(\alpha)d\|E(\alpha)g\|^2=m_b$. In other words, for $B_1=\phi_1(B)$ which belongs to the class of Problem (B), we have $(B_1g, g)=m_b$.

Thus $B_1=\phi_1(B)$ is a solution of Problem (B), too. But one verifies easily that $B_1=\phi_1(B)=1-E(\lambda)$, and so B_1 is a projection.

Therefore replacement of B_0 by B_1 completes the proof.

LEMMA 3.2.3. *Problems* (A) *and* (*B*) *possess a common solution* $E_0=B_0$, *and for the maxima we have* $m_a=m_b$.

The class of the E in Problem (A) is obviously a subclass of the B in Problem (B): It consists of all projections of the latter class. (This is so, because for projections E, $D_M(E)=Tr_M(E)$ (cf. R.O., Lemma 15.3.1).) Now Lemma 3.2.2 states, that a solution of Problem (B) exists, which is a projection and thus belongs to the class of Problem (A). Therefore it is a solution of Problem (A) too, and $m_a=m_b$.

Problems (A) and (B) can be modified, when a projection $G=P_{\mathfrak{N}}\epsilon\mathfrak{M}$ is given, in the following sense: Replace $\mathfrak{H}$, $\boldsymbol{M}$ by $\mathfrak{N}$, $\boldsymbol{M}_{(\mathfrak{N})}$ (cf. R.O., §11.3). Then $D_M(E)$ must be replaced by $D_M(E)/D_M(G)$, in order to conserve the standard normalization and so the normalization of §1.1 requires replacing of $D_{M'}(E')$ by $D_{M'}(E')/D_M(G)$. This α is replaced by $\alpha/D_M(G)$ and so $\alpha\geqq 1$ remains true. Lemma 3.2.3 is modified, in so far as λ is replaced by $\lambda/D_M(G)$ and $\boldsymbol{M}$ in Problems (A) and (B) is to be replaced by $\boldsymbol{M}_{(\mathfrak{N})}$. This means that projections $F\epsilon\boldsymbol{M}$ must be $\leqq G$, and operators $B\epsilon\boldsymbol{M}$ must fulfill $BG=GB=B$.

Combining this and the corresponding changes in Lemma 3.2.1 (observe that $(B_0g, g)=(E_0g, g)=\|E_0g\|^2$), we have

COROLLARY. *Let a projection* $G\epsilon\boldsymbol{M}$ *be given. Replace Problems* (A) *and* (B) *by Problems* (A_G) *and* (B_G), *which arise by imposing these further restrictions: In* (A_G) *the projections* $F\epsilon\boldsymbol{M}$ *are* $F\leqq G$. *In* (B_G) *the operators* $B\epsilon\boldsymbol{M}$ *fulfill* $BG=GB=B$. *Assume* $0\leqq\lambda\leqq D_M(G)$. *Then Problems* ($A_G$) *and* ($B_G$) *possess a common solution* $E_0=B_0$, *and we have for the maxima* $K^{-1}\|E_0g\|^2\leqq m_a=m_b\leqq K\|E_0g\|^2$.

3.3. We hold g fixed, and make the following

DEFINITION 3.3.1. *Let* E, G *be two projections* $\epsilon\boldsymbol{M}$, $E\leqq G$. *We say that* E *reduces* g *with respect to* G *if for every* $A\epsilon\boldsymbol{M}$ *with* $AG=GA=A$

$$(AEg, g) = (EAg, g).$$

If G *may be taken* $=1$, *we say that* E *reduces* g.

LEMMA 3.3.1. *Let* G *be a projection* $\epsilon\boldsymbol{M}$, E_0 *a solution of Problem* (A_G) (cf. the corollary to Lemma 3.2.3). *Then* E_0 *reduces* g *with respect to* G.

If U is unitary, $\epsilon\boldsymbol{M}$, and commutes with G, then $U^{-1}E_0U$ is a projection $\epsilon\boldsymbol{M}$, it is $\leqq U^{-1}GU=G$, and $D_M(U^{-1}E_0U)=D_M(E_0)=\lambda$. So $(U^{-1}E_0Ug, g)$

$\leqq(E_0g, g)$. Now $(U^{-1}E_0Ug, g)=(U^*E_0Ug, g)=(E_0Ug, Ug)$. So $(E_0g, g)-(E_0Ug, Ug)\geqq 0$.

Let A be Hermitian, ϵM, and commute with G. Put for some $\epsilon\gtrless 0$ $\phi_\epsilon(\alpha)=(1+i\epsilon\alpha)/(1-i\epsilon\alpha)$. As $|\phi_\epsilon(\alpha)|=1$, so $U_\epsilon=\phi_\epsilon(A)$ is unitary, and it is ϵM and commutes with G along with A. Further $\phi_\epsilon(\alpha)=1+2i\epsilon\alpha+\epsilon^2\psi_\epsilon(\alpha)$, where $\psi_\epsilon(\alpha)=-\alpha^2/(1-i\epsilon\alpha)$. So $|\psi_\epsilon(\alpha)|\leqq D^2$ if $|\alpha|\leqq D$, and therefore $|||\psi_\epsilon(A)|||\leqq |||A|||^2$ (cf., for instance, (16), p. 113, Theorem 4*). Now

$$\begin{aligned} 0 \leqq (E_0g, g) - (E_0Ug, Ug) &= (E_0g, g) \\ &\quad - (E_0(1 + 2i\epsilon A + \epsilon^2\psi_\epsilon(A))g,\ (1 + 2i\epsilon A + \epsilon^2\psi_\epsilon(A))g) \\ &= 2\epsilon(i(AE_0 - E_0A)g, g) + O(\epsilon^2). \end{aligned}$$

So we have

$$2\epsilon i((AE_0 - E_0A)g, g) + O(\epsilon^2) \geqq 0.$$

As $\epsilon\gtrless 0$, this necessitates $((AE_0-E_0A)g, g)=0$, that is,

$$(AE_0g, g) = (E_0Ag, g).$$

This equation extends from the Hermitian $A\epsilon M$ to all $A\epsilon M$, which commute with G. Put for such an A, $A_1=\frac{1}{2}(A+A^*)$, $A_2=-\frac{1}{2}i(A-A^*)$, then it holds for A_1, A_2 and so for $A=A_1+iA_2$. Thus we have established that E_0 reduces g with respect to G.

LEMMA 3.3.2. *If E reduces g relatively to G and G reduces g, then E reduces g.*

We must show that for every $A\epsilon M$, $(AEg, g)=(EAg, g)$. Now

$$\begin{aligned} (EAg, g) &= (GEAg, g) = (G(EA)g, g) = ((EA)Gg, g) = (EGAGg, g) \\ &= (GAGEg, g) = (GAEg, g) = ((AE)Gg, g) = (AEg, g). \end{aligned}$$

LEMMA 3.3.3. *If $\{E_i\}$ is a sequence of projections $E_i\epsilon M$ each of which reduce g and $E_i\cdot E_j=0$ if $i\neq j$, then $\sum_{i=1}^\infty E_i$ reduces g. If E_1 and E_2 reduce g and $E_1\geqq E_2$ then E_1-E_2 reduces g.*

This is immediate for $A\epsilon M$ implies A is bounded.

LEMMA 3.3.4. *Let $\{E_i\}$ be a sequence with the same properties as in Lemma 3.3.3. Let $E=\sum_{i=1}^\infty E_i$. Then for $A\epsilon M$,*

$$(EAg, g) = (AEg, g) = (EAEg, g) = \sum_{i=1}^{\infty} (E_iAE_ig, g).$$

Also if E reduces g, $EF=0$, $F\epsilon M$, then $(EAFg, g)=(FAEg, g)=0$.

We have $(EAg, g)=(E(EA)g, g)=(EAEg, g)=(E(AE)g, g)=((AE)Eg, g)=(AEg, g)$. Hence we have $(EAg, g)=((\sum_{i=1}^\infty E_i)Ag, g)=\sum_{i=1}^\infty(E_iAg, g)$

$=\sum_{i=1}^{\infty}(E_iAE_ig, g)$. Also $(EAFg, g)=(E(AF)g, g)=((AF)Eg, g)=(A(FE)g, g)$ $=0=((EF)Ag, g)=(E(FA)g, g)=((FA)Eg, g)=(FAEg, g)$.

LEMMA 3.3.5. *Let $\{E_i\}$ be a sequence of mutually orthogonal projections. If $E=\sum_{i=1}^{\infty}E_i$ and A and $E_i \epsilon M$,*

$$Tr_M(EAE) = \sum_{i=1}^{\infty} Tr_M(E_iAE_i).$$

By Property III, (vi), $Tr_M(EAE)=Tr_M(AE\cdot E)=Tr_M(AE)$. Similarly $Tr_M(E_iAE_i)=Tr_M(AE_i)$. By Property III, $(**)$, and (iii) $Tr_M(AE)$ $=Tr_M(\sum_{i=1}^{\infty}AE_i)=\sum_{i=1}^{\infty}Tr_M(AE_i)$. Thus $Tr_M(EAE)=\sum_{i=1}^{\infty}Tr_M(E_iAE_i)$. Note that we have also demonstrated the convergence of $\sum_{i=1}^{\infty}Tr_M(E_iAE_i)$.

LEMMA 3.3.6. *Let $\{E_i\}$ be a sequence of projections each $E_i \geqq {}_p\lambda$ $(\leqq {}_p\lambda)$ for g. Furthermore let us suppose that each E_i reduces g and $E_i\cdot E_j=0$ if $i\neq j$. Then $E=\sum_{i=1}^{\infty}E_i \geqq {}_p\lambda$ $(\leqq {}_p\lambda)$.*

Let A be positive definite, ϵM and such that $EA=AE=A$. Then by Lemma 3.3.4

$$(Ag, g) = (EAg, g) = \sum_{i=1}^{\infty} (E_iAE_ig, g).$$

Now E_iAE_i is positive definite and is unchanged by left- or right-multiplication with E_i. Hence Lemma 1.4.2 yields $(E_iAE_ig, g) \geqq \lambda Tr_M(E_iAE_i)$. Thus by Lemma 3.2.6

$$(Ag, g) = \sum_{i=1}^{\infty} (E_iAE_ig, g) \geqq \lambda \sum_{i=1}^{\infty} Tr_M(E_iAE_i) = \lambda Tr_M(A).$$

Lemma 1.4.2 now implies $E \geqq {}_p\lambda$.

3.4. We still suppose that g is held fixed.

LEMMA 3.4.1. *Let $E\neq 0$ be a projection, ϵM, and such that $a \leqq {}_pE \leqq {}_pb$. Let a λ with $0<\lambda<D_M(E)$ be given. Then there exists a projection $E_0 \epsilon M$ with the following properties:* (i) $E_0 \leqq E$; (ii) $D_M(E_0)=\lambda$; (iii) *E_0 reduces g with respect to E;* and (iv) *there exists a ξ_0 with $a \leqq {}_pE-E_0 \leqq {}_p\xi_0 \leqq {}_pE_0 \leqq {}_pb$.*

Consider the solution E_0 of Problem (A_G) with $G=E$, in the corollary to Lemma 3.2.3. This E_0 is a projection ϵM and fulfills (i), (ii) by definition, and it fulfills (iii) by Lemma 3.3.1. As to (iv), both E_0 and $E-E_0$ are $\geqq {}_pa$ and $\leqq {}_pb$ along with E, by Lemma 1.2.3.

So we must only find a ξ with $E-E_0 \leqq {}_p\xi \leqq {}_pE_0$. Let Γ' be the set of all η's, for which there exists an $F \leqq E_0$ with $F<\eta$; and let Γ'' be the set of all η's for which there exists an $F \leqq E-E_0$ with $F>\eta$. Γ' and Γ'' are open in-

tervals, and clearly every $\eta > b$ belongs to Γ', and every $\eta < a$ belongs to Γ''. So there exists either a ξ_0 such that every η in Γ' is $\geqq \xi_0$ and every η in Γ'' is $\leqq \xi_0$, or Γ' and Γ'' have a common element ξ_1.

If the former holds, then $F \leqq E_0$ implies $F \geqq \eta$ for all $\eta < \xi_0$, and so $F \geqq \xi_0$, that is, $E_0 \geqq {}_p\xi_0$. Similarly $E - E_0 \leqq {}_p\xi_0$. So in this case the rest of (iv) is proved, too.

If the latter holds, then we have even a $\xi_1 - \delta$ in Γ' and $\xi_1 + \delta$ in Γ'' for some suitable $\delta > 0$. So an $F_1 \leqq E_0$ with $F_1 < \xi_1 - \delta$ exists, and an $F_2 \leqq E - E_0$ with $F_2 > \xi_1 + \delta$. By Lemma 1.2 there exist two F_3 and F_4 (both $\neq 0$), with $F_3 \leqq F_1$, $F_4 \leqq F_2$ and $F_3 \leqq {}_p\xi_1 - \delta$, $F_4 \geqq {}_p\xi_1 + \delta$. We may assume that $D_M(F_3) = D_M(F_4)$, since otherwise we may replace F_3, F_4 by two $F_3' \leqq F_3$, $F_4' \leqq F_4$ with $D_M(F_3') = D_M(F_4') = \min(D_M(F_3), D_M(F_4))$ (remembering Lemma 1.2.3).

Now as $F_3 \leqq E_0$, $F_4 \leqq E - E_0$, so $E_0 - F_3 + F_4$, is a projection $\leqq E$, and

$$D_M(E_0 - F_3 + F_4) = D_M(E_0) - D_M(F_3) + D_M(F_4) = D_M(E_0) = \lambda,$$

while

$$\begin{aligned}((E_0 - F_3 + F_4)g, g) &= (E_0 g, g) - (F_3 g, g) + (F_4 g, g) \\ &\geqq m_a - (\xi_1 - \delta)D_M(F_3) + (\xi_1 + \delta)D_M(F_4) \\ &\geqq m_a - (\xi_1 - \delta)D_M(F_3) + (\xi_1 + \delta)D_M(F_3) \\ &= m_a + 2\delta D_M(F_3) > m_a,\end{aligned}$$

contradicting the maximum property of m_a. So this case cannot arise.

The proof is therefore completed.

LEMMA 3.4.2. *We can define for all* $0 \leqq \alpha \leqq 1$ *a family of projections* $E(\alpha)$ *and a function* $\xi(\alpha)$ *with the following properties*:

(i) $E(0) = 0$, $E(1) = 1$, (ii) $\alpha \leqq \beta$ *implies* $E(\alpha) \leqq E(\beta)$,
(iii) $\xi(0) = K$, $\xi(1) = K^{-1}$, (iv) $\alpha \leqq \beta$ *implies* $\xi(\alpha) \geqq \xi(\beta)$,
(v) $D_M(E(\alpha)) = \alpha$, (vi) $E(\alpha)$ *reduces* g,
(vii) $\alpha < \beta$ *implies* $\xi(\beta - 0) \leqq {}_pE(\beta) - E(\alpha) \leqq {}_p\xi(\alpha + 0)$.

Choose a sequence $\rho_1, \rho_2, \cdots$ which lies and is everywhere dense in $0 \leqq \alpha \leqq 1$, with $\rho_1 = 0$, $\rho_2 = 1$. We will define $E(\alpha)$ and $\xi(\alpha)$ for $\alpha = \rho_1, \rho_2, \cdots$ so that they fulfill (i)–(vi) and

(vii)′ $E(\alpha) \geqq {}_p\xi(\alpha)$, $1 - E(\alpha) \leqq {}_p\xi(\alpha)$,

for all $\alpha = \rho_1, \rho_2, \cdots$.

Put first $E(0) = 0$, $E(1) = 1$, $\xi(0) = K$, $\xi(1) = K^{-1}$. Then $\alpha = \rho_1, \rho_2$ are taken care of, and (i)–(vi) as well as (vii)′ hold for $\alpha = \rho_1, \rho_2$. Assume now that for a $j = 3, 4, \cdots$ the $E(\alpha)$, $\xi(\alpha)$ for $\alpha = \rho_1, \cdots, \rho_{j-1}$ are already defined, so that (i)–(vi), (vii)′ hold for $\alpha = \rho_1, \cdots, \rho_{j-1}$, we will now define $E(\rho_j)$, $\xi(\rho_j)$ without violating (i)–(vi), (vii)′.

Consider the ρ_i, $i=1, \cdots, j-1$ with $\rho_i \leqq \rho_j$ (they exist: $\rho_1 = 0 \leqq \rho_j$); let $\rho_{i'}$ be the greatest one. Consider the ρ_i, $i=1, \cdots, j-1$, with $\rho_i \geqq \rho_j$ (they exist: $\rho_2 = 1 \geqq \rho_j$); let $\rho_{i''}$ be the smallest one. As $\rho_{i'}, \rho_{i''} \neq \rho_j$, so we have $\rho_{i'} < \rho_j < \rho_{i''}$. So (ii) gives $E(\rho_{i'}) \leqq E(\rho_{i''})$, and (v) gives $D_M(E(\rho_{i''}) - E(\rho_{i'})) = \rho_{i''} - \rho_i > \rho_j - \rho_{i'} > 0$.

Apply now Lemma 3.4.1 to $E = E(\rho_{i''}) - E(\rho_{i'})$, $\lambda = \rho_j - \rho_{i'}$ with $a = \xi(\rho_{i''})$, $b = \xi(\rho_{i'})$. Owing to (vii)′ and Lemma 1.2.3 we have $E(\rho_{i''}) - E(\rho_{i'}) \leqq E(\rho_{i''})$ and therefore $\geqq_p \xi(\rho_{i''})$, and $E(\rho_{i''}) - E(\rho_{i'}) \leqq 1 - E(\rho_{i'})$ and therefore $\leqq_p \xi(\rho_{i'})$. That is, $a \leqq_p E \leqq_p b$. An $E_0 \leqq E(\rho_{i''}) - E(\rho_{i'})$ and a ξ_0 result, put $E(\rho_j) = E(\rho_{i'}) + E_0$ (this is a projection ϵM) and $\xi(\rho_j) = \xi_0$. Let us now consider (i)–(vi), (vii)′ for $\alpha = \rho_1, \cdots, \rho_{j-1}, \rho_j$.

(i), (iii) are unaffected.

In (ii), (iv) the only new possibilities are $\alpha = \rho_i \leqq \rho_j = \beta$ and $\alpha = \rho_j \leqq \rho_i = \beta$, where $i = 1, \cdots, j-1$. In the first case $\rho_i \leqq \rho_{i'}$, so $E(\alpha) = E(\rho_i) \leqq E(\rho_{i'}) \leqq E(\rho_{i'}) + E_0 = E(\rho_j) = E(\beta)$, and $\xi(\alpha) = \xi(\rho_i) \geqq \xi(\rho_{i'}) = b \geqq \xi_0 = \xi(\rho_j) = \xi(\beta)$. In the second case $\rho_i \geqq \rho_{i''}$, so $E(\beta) = E(\rho_i) \geqq E(\rho_{i''}) = E(\rho_i) + E(\rho_{i''}) - E(\rho_{i'}) \geqq E(\rho_i) + E_0 = E(\rho_j) = E(\alpha)$, and $\xi(\beta) = \xi(\rho_i) \leqq \xi(\rho_{i''}) = a \leqq \xi_0 = \xi(\rho_j) = \xi(\alpha)$. So (ii), (iv) remain true.

In (v) we need only $D_M(E(\rho_j)) = \rho_j$. Now $D_M(E(\rho_j)) = D_M(E(\rho_{i'}) + E_0) = D_M(E(\rho_{i'})) + D_M(E_0) = \rho_{i'} + (\rho_j - \rho_{i'}) = \rho_j$.

In (vi) we need only that $E(\rho_j)$ reduces g. As $E(\rho_j) = E(\rho_{i'}) + E_0$, and $E(\rho_{i'})$ reduces g, we must only show that E_0 reduces g (use Lemma 3.3.3). But E_0 reduces g with respect to $E = E(\rho_{i''}) - E(\rho_{i'})$ by Lemma 3.4.1, (iii), and $E(\rho_{i''}) - E(\rho_{i'})$ reduces g because $E(\rho_{i''})$ and $E(\rho_{i'})$ do (again use Lemma 3.3.3), therefore E_0 reduces g by Lemma 3.3.2. So the property (vi) remains true.

In (vii)′ we need only $E(\rho_j) \geqq_p \xi(\rho_j)$, $1 - E(\rho_j) \leqq_p \xi(\rho_j)$. Now $E(\rho_j) = E(\rho_{i'}) + E_0$, and $E(\rho_{i'}) \geqq_p \xi(\rho_{i'}) \geqq \xi(\rho_j)$, $E_0 \geqq_p \xi_0 = \xi(\rho_j)$ by Lemma 3.4.1, (iv), so $E(\rho_j) \geqq_p \xi(\rho_j)$ by Lemma 3.3.6. On the other hand $1 - E(\rho_j) = 1 - E(\rho_{i''}) + ((E(\rho_{i''}) - E(\rho_{i'})) - E_0)$ and by Lemma 3.4.1, (iv), $1 - E(\rho_{i''}) \leqq_p \xi(\rho_{i''}) \leqq \xi(\rho_j)$, $((E(\rho_{i''}) - E(\rho_{i'})) - E_0) = E - E_0 \leqq_p \xi_0 = \xi(\rho_j)$ so that $1 - E(\rho_j) \leqq_p \xi(\rho_j)$ by Lemma 3.3.6. Therefore (vii)′ remains true.

Thus we have verified (i)–(vi), (vii)′ for $\alpha = \rho_1, \cdots, \rho_{j-1}, \rho_j$. Therefore all $\alpha = \rho_1, \rho_2, \cdots$ are taken care of, and (i)–(vi), (vii)′ hold for all of them. Owing to (ii) and (iv) $\lim_{\rho_i \to \alpha, \rho_i \geqq \alpha} E(\rho_i)$ and $\lim_{\rho_i \to \alpha, \rho_i > \alpha} \xi(\rho_i)$ exist for all α with $0 \leqq \alpha \leqq 1$. If α is equal to a ρ_j, then this limit is meant to denote the value at ρ_j. So we can extend the definitions of $E(\alpha)$ and $\xi(\alpha)$ to all above α by defining

$$E(\alpha) = \lim_{\rho_i \to \alpha, \rho_i \geqq \alpha} E(\rho_i), \qquad \xi(\alpha) = \lim_{\rho_i \to \alpha, \rho_i \geqq \alpha} \xi(\rho_i).$$

The statements (i), (iii) are unaffected by this extension, while (ii), (iv)–(vi) extend by continuity to all $0 \leqq \alpha \leqq 1$.

Let us now consider (vii). Assume $\alpha < \beta$. Choose a sequence ρ_{i_n}, $n = 0, \pm 1, \pm 2, \cdots$, with $\alpha < \cdots < \rho_{i_{-2}} < \rho_{i_{-1}} < \rho_{i_0} < \rho_{i_1} < \rho_{i_2} < \cdots < \beta$ and $\lim_{n\to-\infty} \rho_{i_n} = \alpha$, $\lim_{n\to\infty} \rho_{i_n} = \beta$. We have $\rho_{i_{-1}} > \rho_{i_{-2}} > \cdots > \alpha$, so that $E(\rho_{i_{-1}}) \geqq E(\rho_{i_{-2}}) \geqq \cdots \geqq E(\alpha)$. Therefore $\lim_{n\to-\infty} E(\rho_{i_n})$ exists, and is $\geqq E(\alpha)$. Now we have

$$D_M\left(\lim_{n\to-\infty} E(\rho_{i_n}) - E(\alpha)\right) = \lim_{n\to-\infty} D_M(E(\rho_{i_n})) - D_M(E(\alpha))$$

$$= \lim_{n\to-\infty} \rho_{i_n} - \alpha = \alpha - \alpha = 0, \quad \lim_{n\to-\infty} E(\rho_{i_n}) = E(\alpha).$$

Similarly

$$\lim_{n\to\infty} E(\rho_{i_n}) = E(\beta).$$

Therefore

$$\sum_{m=-\infty}^{\infty} (E(\rho_{i_{m+1}}) - E(\rho_{i_m})) = \lim_{n\to\infty} \sum_{m=-n}^{n-1} (E(\rho_{i_{m+1}}) - E(\rho_{i_m}))$$

$$= \lim_{n\to\infty} (E(\rho_{i_n}) - E(\rho_{i_{-n}})) = E(\beta) - E(\alpha).$$

We have $E(\rho_{i_{m+1}}) - E(\rho_{i_m}) \leqq E(\rho_{i_{m+1}})$, therefore (vii)′ (for $\alpha = \rho_{i_{m+1}}$) and Lemma 1.2.3 give $E(\rho_{i_{m+1}}) - E(\rho_{i_m}) \geqq {}_p\xi(\rho_{i_{m+1}}) \geqq \xi(\beta - 0)$. Then Lemma 3.3.6 gives $E(\beta) - E(\alpha) \geqq {}_p\xi(\beta - 0)$. On the other hand we have $E(\rho_{i_{m+1}}) - E(\rho_{i_m}) \leqq 1 - E(\rho_{i_m})$ therefore (vii)′ (for $\alpha = \rho_{i_m}$) and Lemma 1.2.3 give $E(\rho_{i_{m+1}}) - E(\rho_{i_m}) \leqq {}_p\xi(\rho_{i_m}) \leqq \xi(\alpha + 0)$. Then Lemma 3.3.6 gives $E(\beta) - E(\alpha) \leqq {}_p\xi(\alpha + 0)$. Thus we have proved

$$\xi(\beta - 0) \leqq {}_pE(\beta) - E(\alpha) \leqq {}_p\xi(\alpha + 0),$$

that is, (vii).

We have established (i)–(vii) for all $0 \leqq \alpha \leqq 1$, and so the proof is completed.

Lemma 3.4.3. *Let α_i, $i = 0, 1, \cdots, n$ be a set of numbers such that $0 = \alpha_0 \leqq \alpha_1 \leqq \cdots \leqq \alpha_{n-1} \leqq \alpha_n = 1$ and let λ_i, $i = 1, \cdots, n$ be a set of positive real numbers. Put (with the $E(\alpha)$, $\xi(\alpha)$ of Lemma 3.4.2)*

$$g' = \sum_{i=1}^{n} \lambda_i(E(\alpha_i) - E(\alpha_{i-1}))g,$$

$$C_1 = \max_{i=1,\cdots,n} \lambda_i^2 \xi(\alpha_{i-1} + 0); \qquad C_2 = \min_{i=1,\cdots,n} \lambda_i^2 \xi(\alpha_i - 0).$$

Then for all definite $A \epsilon M$

$$C_1 Tr_M(A) \geqq (Ag', g') \geqq C_2 Tr_M(A).$$

We have

$$(Ag', g') = \left(A\sum_{i=1}^{n} \lambda_i(E(\alpha_i) - E(\alpha_{i-1}))g,\ \sum_{j=1}^{n} \lambda_j(E(\alpha_j) - E(\alpha_{j-1}))g\right)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j(A(E(\alpha_i) - E(\alpha_{i-1}))g,\ (E(\alpha_j) - E(\alpha_{j-1}))g)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j((E(\alpha_j) - E(\alpha_{j-1}))A(E(\alpha_i) - E(\alpha_{i-1}))g,\ g).$$

As every $E(\alpha_k)-E(\alpha_{k-1})$ reduces g (by Lemma 3.4.2, (vi), and Lemma 3.3.3) and as $(E(\alpha_j)-E(\alpha_{j-1}))\cdot(E(\alpha_i)-E(\alpha_{i-1}))=0$ for $j\neq i$, Lemma 3.3.4 permits us to drop all terms with $i\neq j$ from the above sum $\sum_{i,j=1}^{n}$. Therefore it becomes $\sum_{i=1}^{n}\lambda_i^2((E(\alpha_i)-E(\alpha_{i-1}))A(E(\alpha_i)-E(\alpha_{i-1}))g,\ g)$.

Now $E(\alpha_i)-E(\alpha_{i-1})\leqq{}_p\xi(\alpha_{i-1}+0)$ by Lemma 3.4.2, (vii). Therefore it is evident from Lemma 1.4.1 that $((E(\alpha_i)-E(\alpha_{i-1}))A(E(\alpha_i)-E(\alpha_{i-1}))g,\ g)$ $\leqq\xi(\alpha_{i-1}+0)Tr_M((E(\alpha_i)-E(\alpha_{i-1}))A(E(\alpha_i)-E(\alpha_{i-1})))$. So

$$(Ag', g') \leqq \sum_{i=1}^{n} \lambda_i^2\xi(\alpha_{i-1} + 0)Tr_M((E(\alpha_i) - E(\alpha_{i-1}))A(E(\alpha_i) - E(\alpha_{i-1})))$$
$$\leqq C_1\sum_{i=1}^{n} Tr_M((E(\alpha_i) - E(\alpha_{i-1}))A(E(\alpha_i) - E(\alpha_{i-1})))$$
$$\leqq C_1Tr_M(A)$$

(use Lemma 3.3.5). Similarly $E(\alpha_i)-E(\alpha_{i-1})\geqq{}_p\xi(\alpha_i-0)$ by Lemma 3.4.2, (vii). Therefore Lemma 1.4.1 gives $((E(\alpha_i)-E(\alpha_{i-1}))A(E(\alpha_i)-E(\alpha_{i-1}))g, g)$ $\geqq\xi(\alpha_i-0)Tr_M((E(\alpha_i)-E(\alpha_{i-1}))A(E(\alpha_i)-E(\alpha_{i-1})))$. So

$$(Ag', g') \geqq \sum_{i=1}^{n} \lambda_i^2\xi(\alpha_i - 0)Tr_M((E(\alpha_i) - E(\alpha_{i-1}))A(E(\alpha_i) - E(\alpha_{i-1})))$$
$$\geqq C_2\sum_{i=1}^{n} Tr_M((E(\alpha_i) - E(\alpha_{i-1}))A(E(\alpha_i) - E(\alpha_{i-1}))) = C_2Tr_M(A)$$

(use Lemma 3.3.5). Thus the proof is completed.

3.5. We can now take the decisive step.

Lemma 3.5.1. *Let g and K be as in Theorem* I. *Let K' be any number such that $1<K'\leqq K$. Then there is a g' which is related to K' as g is to K in Theorem* I, *and such that*

$$\|g' - g\| \leqq (K - 1)K^{1/2}.$$

Let n be the smallest positive integer with $(K')^n\geqq K$. Put $\theta=K^{1/n}$, so

$1<\theta\leqq K'$. Choose the α_i, $i=0, 1, \cdots, n$ with $0=\alpha_0\leqq\alpha_1\leqq\cdots\leqq\alpha_{n-1}\leqq\alpha_n=1$, so that $\xi(\alpha_i+0)\leqq\theta^{n-2i}\leqq\xi(\alpha_i-0)$. (For $i=0$, $\theta^{n-2i}=\theta^n=K$; for $i=n$, $\theta^{n-2i}=\theta^{-n}=K^{-1}$.) Put $\lambda_i=\theta^{i-(n+1)/2}$. Now apply Lemma 3.4.3. We have $\lambda_i^2\xi(\alpha_{i-1}+0)\leqq\theta^{2i-(n+1)}\theta^{n-2(i-1)}=\theta\leqq K'$ and $\lambda_i^2\xi(\alpha_i-0)\geqq\theta^{2i-(n+1)}\cdot\theta^{n-2i}\geqq\theta^{-1}\geqq(K')^{-1}$, so $C_1\leqq K'$ and $C_2\geqq(K')^{-1}$. Therefore the g' of that lemma gives for every definite $A\epsilon \boldsymbol{M}$: $K'Tr_{\boldsymbol{M}}(A)\geqq(Ag', g')\geqq(K')^{-1}Tr_{\boldsymbol{M}}(A)$; that is, $K'(Ag', g')\geqq Tr_{\boldsymbol{M}}(A)\geqq(K')^{-1}(Ag', g')$. So it is related to K' as g and K are in Theorem I.

Compute now $\|g'-g\|$. The projections $E(\alpha_i)-E(\alpha_{i-1})$, $i=1, \cdots, n$, are mutually orthogonal, therefore the $(E(\alpha_i)-E(\alpha_{i-1}))g$, $i=1, \cdots, n$ are mutually orthogonal too. Now we have

$$\begin{aligned}\|g'-g\|^2 &= \left\|\sum_{i=1}^{n}\lambda_i(E(\alpha_i)-E(\alpha_{i-1}))g-\sum_{i=1}^{n}(E(\alpha_i)-E(\alpha_{i-1}))g\right\|^2\\ &= \left\|\sum_{i=1}^{n}(\lambda_i-1)(E(\alpha_i)-E(\alpha_{i-1}))g\right\|^2\\ &= \sum_{i=1}^{n}(\lambda_i-1)^2\|(E(\alpha_i)-E(\alpha_{i-1}))g\|^2\\ &\leqq \max_{i=1,\cdots,n}(\lambda_i-1)^2\sum_{i=1}^{n}\|(E(\alpha_i)-E(\alpha_{i-1}))g\|^2\\ &= \max_{i=1,\cdots,n}(\lambda_i-1)^2\cdot\|g\|^2.\end{aligned}$$

Now clearly

$$\max_{i=1,\cdots,n}(\lambda_i-1)^2=(\lambda_n-1)^2=(\theta^{(n+1)/2}-1)^2\leqq(\theta^n-1)^2=(K-1)^2$$

so that $\|g'-g\|\leqq(K-1)\|g\|$.

But $A=1$ in Theorem I gives

$$\|g\|^2=(g, g)\leqq KTr_{\boldsymbol{M}}(1)=K, \quad \|g\|\leqq K^{1/2}.$$

Therefore $\|g'-g\|\leqq(K-1)K^{1/2}$. Thus the proof is completed.

Let $K_i=1+2^{-(i+2)}$, $i=0, 1, 2, \cdots$. We define inductively a sequence of elements g_i, $i=0, 1, 2, \cdots$. Let g_0 be related to K_0 as g is to k in Theorem I. Suppose g_i has been defined and is related to K_i as in Theorem I. Then in Lemma 3.5.1 let $g=g_i$, $K=K_i$, $K'=K_{i+1}$. We then define g_{i+1} as g'. Then g_{i+1} and K_{i+1} are related as g and K in Theorem I and we also have

$$\|g_{i+1}-g_i\|\leqq(K_i-1)K_i^{1/2}=\frac{1}{2^{i+2}}K_i^{1/2}\leqq\frac{1}{2^{i+1}}\cdot$$

Then for $p>0$

$$\|g_{n+p} - g_n\| = \left\| \sum_{i=n}^{n+p-1} (g_{i+1} - g_i) \right\| \leqq \sum_{i=n}^{n+p-1} \|g_{i+1} - g_i\| \leqq \sum_{i=n}^{n+p-1} \frac{1}{2^{i+1}}$$
$$\leqq \sum_{i=n}^{\infty} \frac{1}{2^{i+1}} = \frac{1}{2^n}.$$

This implies that the g_i's satisfy the Cauchy condition. Let g be their limit, and let A again be definite and $\epsilon \boldsymbol{M}$. Then, since A is bounded,

$$(Ag, g) = \lim_{i \to \infty} K_i(Ag_i, g_i) \geqq Tr_M(A) \geqq \lim_{i \to \infty} K_i^{-1}(Ag_i, g_i) = (Ag, g).$$

Thus, if $A \epsilon \boldsymbol{M}$ is definite, $(Ag, g) = Tr_M(A)$.

For $A \epsilon \boldsymbol{M}$, Hermitian, we have as in §2.1, under Property II, $A = A_1 - A_2$, A_1 and A_2, $\epsilon \boldsymbol{M}$ and positive definite. Property I of §2.1 then yields $(Ag, g) = Tr_M(A)$. If A is merely $\epsilon \boldsymbol{M}$, then Definition 2.2.1 now implies that $(Ag, g) = Tr_M(A)$. We have thus demonstrated

THEOREM II. *Let* $\boldsymbol{M}$ *be a factor in case* II, *with* $\alpha \geqq 1$. (*In the normalization of* §1.1. *In the normalization of R.O., Theorem* VIII *this means* $\boldsymbol{M}$, $\boldsymbol{M}'$ *are either in case* II_1, II_∞ *or in case* II_1, II_1 *with* $C \leqq 1$, *standard normalization.*) *Then there exists a* $g \epsilon \mathfrak{H}$ *such that*

$$(Ag, g) = Tr_M(A).$$

We now drop the restriction that $\alpha \geqq 1$. Let $\boldsymbol{M}$, $\boldsymbol{M}'$ be in case II_1, II_1, $\alpha < 1$, and let m be an integer such that $m\alpha \geqq 1$ or in the standard normalization of R.O., Theorem X, $C/m \leqq 1$. Let $E_1, E_2, \cdots, E_m$ be m projections each of which is $\epsilon \boldsymbol{M}$, $D_M(E_i) = 1/m$, $E_iE_j = 0$ if $i \neq j$. Let the range of E_i be $\mathfrak{M}_i$. Let us recall R.O., §11.3 using $\mathfrak{M}_i$ for $\mathfrak{M}$.

Consider the $\boldsymbol{M}_{(\mathfrak{M}_i)} = \boldsymbol{M}_i$ and $\boldsymbol{M}'_{(\mathfrak{M}_i)} = \boldsymbol{M}'_i$ of Definition 11.3.1. By Lemma 11.3.2, these constitute a factorization in $\mathfrak{M}_i$. Now we can take for $A \epsilon \boldsymbol{M}$ and such that $E_iA = AE_i = A$, $D_{M_i}(A_{E_i}) = D_M(A)$ and for $A' \epsilon \boldsymbol{M}'$, $D_{M'_i}(A_{E_i}) = D_{M'}(A')$, D_{M_i} and $D_{M'_i}$ are dimension functions for $\boldsymbol{M}_i$ and $\boldsymbol{M}'_i$ respectively (Lemma 11.3.6).

The ranges of $D_{M(\mathfrak{M})}$ and $D_{M'(\mathfrak{M})}$ are by Lemma 11.3.7, the intervals $(0, 1/m)$, $(0, \alpha)$ respectively. So we can obtain the standard normalization of R.O., Theorem X by multiplying by m and $1/\alpha$ respectively. But in this latter normalization, by Lemma 11.4.3, $C = (D_M(E_i)/D_{M'}(1))C_0 = C_0/m \leqq 1$, where D_M and $D_{M'}$ refer to the standard normalization for $\boldsymbol{M}$ and $\boldsymbol{M}'$ as in R.O., Theorem X.

By Theorem II above this implies that there is a $g_i^0 \epsilon \mathfrak{M}_i$ such that, if $A_0 \epsilon \boldsymbol{M}_i$, then $Tr_{M'}(A_0) = (A_0 g_i^0, g_i^0)$. But now if $A \epsilon \boldsymbol{M}$ is such that $E_iA = AE_i = A$, let $A_0 = A_{E_i}$. From the above it will be seen that $Tr_M(A)$

$= (1/m)Tr_{M_i}(A_0) = (1/m)(A_0 g_i^0, g_i^0) = (1/m)(Ag_i^0, g_i^0) = (Ag_i, g_i)$, if $g_i = (1/m)^{1/2} g_i^0$.

Now if $A \epsilon M$, we have by Lemma 3.2.6

$$Tr_M(A) = \sum_{i=1}^{m} Tr_M(E_i A E_i) = \sum_{i=1}^{m} (E_i A E_i g_i, g_i)$$

$$= \sum_{i=1}^{m} (A E_i g_i, E_i g_i) = \sum_{i=1}^{m} (A g_i, g_i).$$

If M, M' is in case I_m, $\mathrm{I}_{m'}$, $m < \infty$, we may proceed as follows. By R.O., Lemma 8.6.1, $\mathfrak{H} = E_m \otimes \mathfrak{H}_2$, where E_m is an m-dimensional euclidean space. Let $\phi_1, \cdots, \phi_m$ be a complete orthonormal set in E_m and $f \epsilon \mathfrak{H}_2$ with $\|f\| = 1$. Then if we let $g_i = \phi_i \otimes f$, we easily verify that the above formula for Tr_M holds. Thus we have

THEOREM III. *If M is a factor in a finite case, then there exists a finite number of elements $g_i \epsilon \mathfrak{H}$ such that*

$$Tr_M(A) = \sum_{i=1}^{m} (A g_i, g_i).$$

Suppose that $A \epsilon M$ is held fixed. Then consider the weak neighborhood $U = U(A; g_1, \cdots, g_m; g_1, \cdots, g_m; \epsilon/m)$. $X \epsilon U$ implies $|((X - A)g_i, g_i)| < \epsilon/m$ for $i = 1, \cdots, m$. Hence if X is also ϵM, Theorem III implies $|Tr_M(X) - Tr_M(A)| < \epsilon$. This proves

THEOREM IV. *If M is a factor in a finite case, $Tr_M(A)$ is weakly continuous.*

CHAPTER IV. THE ISOMORPHISM OF M, M' AND $\mathfrak{H}$ (FOR $\alpha = 1$)

4.1. We assume that M, M' is in case II_1, II_1 and $\alpha = 1$ or what is the same thing that $C = 1$ and we have the standard normalization. We know by the discussion of R.O., §§11.3 and 11.4 how all other cases II can be reduced to this case.

As $\alpha = 1$, Theorem II holds. We define

DEFINITION 4.1.1. *A g which satisfies Theorem* II *is uniformly distributed with respect to M; abbreviated g u.d.r. M.*

LEMMA 4.1.1. *If g is u.d.r. M, then* (i) $\|g\| = 1$, (ii) $E_g^{M'} = 1$, (iii) $E_g^{M} = 1$ (cf. R.O., Definition 5.1.1).

To prove (i), in Theorem II take A as 1. Then $1 = Tr_M(1) = (g, g) = \|g\|^2$.

To prove (ii), in Theorem II take A as $E_g^{M'}$. Then

$$1 = \|g\|^2 = (E_g^{M'} g, g) = Tr_M(E_g^{M'}) = D_M(E_g^{M'})$$

or $D_M(E_g^{M'}) = 1$ which in our normalization implies $E_g^{M'} = 1$.

Consider (iii). Since we have the standard normalization and $C=1$, $D_{M'}(E_g^M)=D_M(E_g^{M'})=1$. This implies $E_g^M=1$.

DEFINITION 4.1.2. *Let g be $\epsilon\mathfrak{H}$. Let $Q_g(\boldsymbol{M})$ consist of all operators $Z\epsilon U(\boldsymbol{M})$, i.e., Z is linear closed $\eta\boldsymbol{M}$ and with a dense domain* (cf. R.O., Definition 4.2.1), *for which Zg exists.*

Now consider any $f\epsilon\mathfrak{H}$ such that $E_f^{M'}=E_f^M=1$. By R.O., Lemma 9.2.1, if $h\epsilon\mathfrak{H}$, then $h=XYf$, where X and Y are $\epsilon U(\boldsymbol{M})$. But if $\boldsymbol{M}$ is in case II_1 by R.O., Theorem XV, $Z=[XY]$ is $\epsilon U(\boldsymbol{M})$, and, since $Z\supseteq XY$, Zf exists, and we also have $h=XYf=Zf$; i.e., every $h\epsilon\mathfrak{H}$ is in the form Zf, where Z is $\epsilon U(\boldsymbol{M})$.

Furthermore this Z is uniquely determined by h. For suppose Z_1 and $Z_2\epsilon U(\boldsymbol{M})$ are such that $Z_1f=Z_2f$ or $(Z_1-Z_2)f=0$. Then $[Z_1-Z_2]$ is by R.O., Theorem XV, $\epsilon U(\boldsymbol{M})$ and since $[Z_1-Z_2]\supseteq Z_1-Z_2[Z_1-Z_2]f$ exists and is zero. Let A' be $\epsilon\boldsymbol{M}'$. Then since $A'[Z_1-Z_2]\subseteq[Z_1-Z_2]A'$, $[Z_1-Z_2]A'f=A'[Z_1-Z_2]f=0$. Thus $[Z_1-Z_2]$ is zero on the set of $A'f$, $A'\epsilon\boldsymbol{M}$. But since $E_f^{M'}=1$, this latter set is everywhere dense in $\mathfrak{H}$. Thus $[Z_1-Z_2]$ is zero on a dense set and since it is closed it must be zero. R.O., Theorem XV, then yields

$$Z_1=[Z_1-0]=[Z_1-[Z_1-Z_2]]=[[Z_1-Z_1]+Z_2]=[0+Z_2]=Z_2.$$

So we have proved

LEMMA 4.1.2. *If f is such that $E_f^{M'}=E_f^M=1$ (in particular, if f is u.d.r. $\boldsymbol{M}$), then $h=Zf$ defines a one-to-one correspondence of $\mathfrak{H}$ and the set of all $Z\epsilon U(\boldsymbol{M})$ for which Zf exists (i.e., between $\mathfrak{H}$ and $Q_f(\boldsymbol{M})$).*

Since our assumptions on f are symmetric in $\boldsymbol{M}$ and $\boldsymbol{M}'$, we also have

LEMMA 4.1.3. *Let f be as in Lemma 4.1.2. Then $h=Z'f$ defines a one-to-one correspondence of $\mathfrak{H}$ and the set of all $Z'\epsilon U(\boldsymbol{M}')$ for which $Z'f$ exists (i.e., between $\mathfrak{H}$ and $Q_f(\boldsymbol{M}')$).*

Thus for $h\epsilon\mathfrak{H}$ we have three correspondences defined by the equations $Zf=h=Z'f$, if f is u.d.r. $\boldsymbol{M}$: (1) the correspondence $\mathfrak{Z}_M$ of $\mathfrak{H}$ and $Q_f(\boldsymbol{M})$, (2) the correspondence $\mathfrak{Z}_{M'}$ of $\mathfrak{H}$ and $Q(\boldsymbol{M}')$, and (3) the correspondence $\mathfrak{Z}_{M,M'}$ of $Q_f(\boldsymbol{M})$ and $Q_f(\boldsymbol{M}')$.

4.2. We now investigate these three correspondences. We know that they are one-to-one. They are obviously linear and of course along with $\mathfrak{H}$ the sets $Q_f(\boldsymbol{M})$ and $Q_f(\boldsymbol{M}')$ are linear. But inasmuch as $\mathfrak{Z}_{M,M'}$ is a correspondence of operators, we would like to know the algebraic properties of this correspondence as well as certain properties of $Q_f(\boldsymbol{M})$ and $Q_f(\boldsymbol{M}')$. In particular we would like to know:

(i) To what extent can the operation $[XY]$ be performed in $Q_f(\boldsymbol{M})$ (or $Q_f(\boldsymbol{M}')$)?

(ii) To what extent can the operation X^* be performed in $Q_f(\boldsymbol{M})$ (or $Q_f(\boldsymbol{M}')$)?

(iii) Is the property of being bounded invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$; i.e., since $\boldsymbol{M} \subset Q_f(\boldsymbol{M})$, $\boldsymbol{M}' \subset Q_f(\boldsymbol{M}')$, is $\boldsymbol{M}$ mapped on $\boldsymbol{M}'$ by $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$?

(iv) Are the properties of being Hermitian, definite, or unitary invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$?

LEMMA 4.2.1. *If f is u.d.r. $\boldsymbol{M}$, then every Uf, $U \epsilon \boldsymbol{M}$ and unitary, is u.d.r. $\boldsymbol{M}$ also.*

If $A \epsilon \boldsymbol{M}$, then

$$(AUf, Uf) = (U^{-1}AUf, f) = Tr_{\boldsymbol{M}}(U^{-1}AU) = Tr_{\boldsymbol{M}}(A).$$

LEMMA 4.2.2. *If f and g are each u.d.r. $\boldsymbol{M}$, then there exists a $U' \epsilon \boldsymbol{M}'$ and unitary with $g = U'f$.*

If $h = Af$, $A \epsilon \boldsymbol{M}$, define h^* as Ag. The correspondence $h \rightleftarrows h^*$ is linear, and, since

$$\begin{aligned} \|h^*\|^2 &= \|Ag\|^2 = (Ag, Ag) = (A^*Ag, g) = Tr_{\boldsymbol{M}}(A^*A) = (A^*Af, f) \\ &= (Af, Af) = \|Af\|^2 = \|h\|^2, \end{aligned}$$

it is isometric. Thus the equation $Wh = h^*$ defines a linear isometric and hence one-valued operator W. Since $E_f^{\boldsymbol{M}} = E_g^{\boldsymbol{M}} = 1$, both domain and range of W are everywhere dense and hence its closure $\tilde{W}$ is unitary.

Now if h is in the domain of W, $h = Af$, $A \epsilon \boldsymbol{M}$. Then, for every $B \epsilon \boldsymbol{M}$, $BWh = Bh^* = BAg = WBAf = WBh$ or $BW \subseteq WB$. This implies that $[BW] \subseteq [WB]$. But $B\tilde{W} = \tilde{B}\tilde{W} \subseteq [BW]$, while since B is bounded $[WB] = \tilde{W}B$. Hence $B\tilde{W} \subseteq \tilde{W}B$. If in particular B is unitary, R.O., Lemma 4.2.2 now yields that $\tilde{W}$ is $\eta \boldsymbol{M}'$ and R.O., Lemma 4.2.1 implies that $\tilde{W}$ is $\epsilon \boldsymbol{M}'$.

Letting $A = 1$ in the definition of W, we get $g = Wf = \tilde{W}f$.

LEMMA 4:2.3. *If f and g are u.d.r. $\boldsymbol{M}$, then there exists a $U \epsilon \boldsymbol{M}$ and unitary such that $g = Uf$.*

By Lemma 4.1.2, $g = Zf$ where Z is $\epsilon U(\boldsymbol{M})$. Using the canonical decomposition for Z, $Z = BW$, where B is self-adjoint and definite and W is partially isometric and as indicated in the proof of Lemma 1.4.3 may be taken as unitary; that is, $g = BWf$, where W is unitary and B definite and self-adjoint. We must show that $B = 1$.

By Lemma 4.2.1, $f_0 = Wf$ is u.d.r. $\boldsymbol{M}$. Hence we must show that if f_0 and g are u.d.r. $\boldsymbol{M}$ and $g = Bf_0$, where B is definite and self-adjoint, then $B = 1$.

Let $E(\lambda)$ be the resolution of the identity corresponding to B. If $E(\lambda)=0$ for $\lambda<1$ and $E(\lambda)=1$ for $\lambda>1$, then $B=1$. Thus if B is not 1, either there exists a $\lambda_1<1$ for which $E(\lambda_1)\neq 0$ or a $\lambda_2>1$ for which $E(\lambda_2)\neq 1$. In the first case we have $\|E(\lambda_1)f_0\|^2=(E(\lambda_1)f_0,\ E(\lambda_1)f_0)=(E(\lambda_1)f_0,\ f_0)=Tr_M(E(\lambda_1)) = (E(\lambda_1)g,\ g) = \|E(\lambda_1)g\|^2 = \|E(\lambda_1)Bf_0\|^2 = \|BE(\lambda_1)f_0\|^2 \leqq \lambda_1^2\|E(\lambda_1)f_0\|^2$ or $(1-\lambda_1^2)\|E(\lambda_1)f_0\|^2\leqq 0$. This implies $0=\|E(\lambda_1)f_0\|^2=(E(\lambda_1)f_0,f_0)=Tr_M(E(\lambda_1)) =D_M(E(\lambda_1))$ or $E(\lambda_1)=0$, a contradiction. In the second case a similar type of calculation yields that $\|(1-E(\lambda_2))f_0\|^2=\|B(1-E(\lambda_2))f_0\|^2\geqq\lambda_2^2\|(1-E(\lambda_2))f_0\|^2$ which again implies that $1-E(\lambda_2)=0$, a contradiction. Thus B must be 1 and our lemma is proved.

LEMMA 4.2.4. *If f is u.d.r. $\boldsymbol{M}$, and $U'\epsilon\boldsymbol{M}'$ and unitary, then $U'f$ is u.d.r. $\boldsymbol{M}$.*

If A is $\epsilon\boldsymbol{M}$,

$$(AU'f,\ U'f) = (U'^{-1}AU'f,\ f) = (AU'^{-1}U'f,\ f) = (Af,\ f) = Tr_M(A).$$

LEMMA 4.2.5. *Unitary operators correspond to unitary operators under $\mathfrak{Z}_{M,M'}$; i.e., if f is u.d.r. $\boldsymbol{M}$, $U\epsilon\boldsymbol{M}$, $U'\epsilon\boldsymbol{M}'$ and $Uf=U'f$, then if either U or U' is unitary the other is too.*

Let $U\epsilon\boldsymbol{M}$ be unitary, then by Lemma 4.2.1 Uf is u.d.r. $\boldsymbol{M}$. Then Lemma 4.2.2 yields that there is a unitary $U'\epsilon\boldsymbol{M}$, such that $U'f=Uf$. Thus $U\sim U'$ under $\mathfrak{Z}_{M,M'}$. Similarly Lemmas 4.2.4 and 4.2.3 yield that to every unitary $U'\epsilon\boldsymbol{M}$, there exists a unitary $U\epsilon\boldsymbol{M}$ such that $U'f=Uf$.

LEMMA 4.2.6. *f u.d.r. $\boldsymbol{M}$ is equivalent to f u.d.r. $\boldsymbol{M}'$.*

Let f be u.d.r. $\boldsymbol{M}$ and $A'\epsilon\boldsymbol{M}$. Then consider $T_0(A')=(A'f,f)$. It is linear in A'. By Lemma 4.1.1, (i) $T_0(1)=1$. If A' is definite, then $T_0(A')=(A'f_0,f_0)\geqq 0$. Furthermore we have $T_0(A'^*)=(A'^*f,\ f)=(f,\ A'f)=\overline{(A'f,\ f)}=\overline{T_0(A)}$. Therefore $T_0(A')$ satisfies (i)–(iv) and (v) of Property III with $\boldsymbol{M}'$ instead of $\boldsymbol{M}$.

We would like to verify next (vi)$'$ of the same property for $T_0(A')$. By Lemma 4.2.5 we have for any unitary $U'\epsilon\boldsymbol{M}$

$$\begin{aligned}T_0(U'^{-1}A'U') &= (U'^{-1}A'U'f,\ f) = (A'U'f,\ U'f) = (A'Uf,\ Uf)\\ &= (U^{-1}A'Uf,\ f) = (A'(U^{-1}U)f,\ f) = (A'f,\ f) = T_0(A').\end{aligned}$$

Thus (vi)$'$ holds. Property IV of §2.2 now implies that $T_0(A')=Tr_{M'}(A')$ and hence f is u.d.r. $\boldsymbol{M}'$.

Replacing $\boldsymbol{M}$ by $\boldsymbol{M}'$ in the above argument yields the converse.

We note that the above argument also proves

LEMMA 4.2.7. *If $f\epsilon\mathfrak{H}$ is such that* (i) $\|f\|=1$ *and* (ii) *if to every unitary $U\epsilon\boldsymbol{M}$ there exists a unitary $U'\epsilon\boldsymbol{M}'$ such that $Uf=U'f$, then f is u.d.r. $\boldsymbol{M}$.*

THEOREM V. *Let $f \epsilon \mathfrak{H}$ be such that $\|f\|=1$ and let $\boldsymbol{M}$, $\boldsymbol{M}'$ be in case* II_1, II_1 *with $C=1$. Then the following statements are equivalent:*

(α) *f is u.d.r. $\boldsymbol{M}$.*

(β) *f is u.d.r. $\boldsymbol{M}'$.*

(γ) *If U is unitary and $\epsilon \boldsymbol{M}$, then there exists a unitary $U' \epsilon \boldsymbol{M}'$ such that $Uf=U'f$.*

(δ) *If U' is unitary and $\epsilon \boldsymbol{M}'$, then there exists a unitary $U \epsilon \boldsymbol{M}$ such that $U'f=Uf$.*

Furthermore either (α), (β), (γ), *or* (δ) *implies $E_f^{\boldsymbol{M}}=E_f^{\boldsymbol{M}'}=1$.*

(α) is equivalent to (β) by Lemma 4.2.6. (α) and (γ) are equivalent by Lemmas 4.2.5 and 4.2.7. Replacing $\boldsymbol{M}$ by $\boldsymbol{M}'$, the last mentioned lemmas also show (β) and (δ) to be equivalent. This and Lemma 4.1.1 prove the last statement of the theorem.

In view of Theorem V, we will say that f is uniformly distributed (abbreviated u.d.) if f is u.d.r. $\boldsymbol{M}$.

COROLLARY. *If f is u.d., $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$ maps $\boldsymbol{M}(\subseteq Q_f(\boldsymbol{M}))$ on $\boldsymbol{M}'(\subseteq Q_f(\boldsymbol{M}'))$; that is, the property of being bounded is invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$.*

By Theorem V, the property of being unitary is invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$. This and the linearity of $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$ imply together that the property of having the form $A=\sum_{\nu=1}^{n} a_\nu U_\nu$ ($n=1, 2, \cdots$; a_ν complex, U_ν unitary and in $\boldsymbol{M}$ or $\boldsymbol{M}'$ along with A) is invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$.

We will show that $A \epsilon Q_f(\boldsymbol{M})$ (or $Q_f(\boldsymbol{M}')$) is bounded if and only if it is in this form. Now if A is in this form, it is obviously bounded so we must only show that every bounded A is in this form. Since $A=A_1+iA_2$, A_1 and A_2 Hermitian, it will be sufficient if this is shown for Hermitian A's. Since if A is in this form, cA is also, we may assume that the bound of A is $\leqq 1$. Then $1-A^2$ is definite and $(1-A^2)^{1/2}$ exists. Letting $U=A+i(1-A^2)^{1/2}$, then $U^*=A-i(1-A^2)^{1/2}$ and U is unitary. Furthermore $A=\frac{1}{2}(U+U^*)$.

Now since the form $\sum_{\nu=1}^{n} a_\nu U_\nu$ is invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$, boundedness must be also.

THEOREM VI. *Assume that f_0 is u.d. Then $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$ is an anti-isomorphism of $\boldsymbol{M}$ and $\boldsymbol{M}'$. That is, if $A \sim A'$, $B \sim B'$, then* (i) $\alpha A \sim \alpha A'$; (ii) $A^* \sim A'^*$; (iii) $A+B \sim A'+B'$; (iv) $AB \sim B'A'$; (v) *if A (or A') is Hermitian, then A' (or A) is also.*

(i) and (iii) are obvious.

Consider (iv). We have $ABf_0=AB'f_0=B'Af_0=B'A'f_0$, so that $AB \sim B'A'$. To prove (v), let U be unitary, $U \sim U'$. Then U' is unitary also by Theorem V. Similarly inasmuch as U^{-1} is unitary, if $U^{-1} \sim V'$, V' is unitary. As

$UU^{-1}=U^{-1}U=1$ by (iv) we have $U'V'=V'U'=1$, and hence $V'=(U')^{-1}$. So $U^{-1}\sim(U')^{-1}$ and $\alpha(U+U^{-1})\sim\alpha(U'+(U')^{-1})$ by (iii). If α is >0, the right side is Hermitian and the left side is any Hermitian element of $\boldsymbol{M}$ (cf. the proof of the corollary of Theorem V). This proves (v).

We now prove (ii). If A is $\epsilon\boldsymbol{M}$, then $A=B+iC$, $A^*=B-iC$, B, C are Hermitian and $\epsilon\boldsymbol{M}$. If $A\sim A'$, $B\sim B'$, $C\sim C'$, then B', C' are Hermitian and $A\sim B'+iC'=A'$, $A^*\sim B'-iC'=A'^*$ proving (ii).

COROLLARY. *The properties of being Hermitian, being definite, being a projection, as well as the numerical quantities* $|||A|||$ *(the bound of* A*) and* $Tr_{\boldsymbol{M}}(A)$ *(or* $Tr_{\boldsymbol{M}'}(A')$*), all within* $\boldsymbol{M}$ *(or* $\boldsymbol{M}'$*), are invariant under* $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$.

A Hermitian means $A=A^*$; A a projection means $A=A^*=A^2$; A definite means that there exists an Hermitian B with $A=B^2$. So all these properties are invariant under $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$. 1 is invariant ($1\epsilon\boldsymbol{M}$ and $1\epsilon\boldsymbol{M}'$, $1f_0=1f_0$). $|||A|||$ is the smallest $\alpha\geqq 0$ such that $\alpha^2\cdot 1-A^*A$ is definite, so $|||A|||$ is invariant. If $A\sim A'$, then $Af_0=A'f_0$ and $Tr_{\boldsymbol{M}}(A)=(Af_0, f_0)=(A'f_0, f_0)=Tr_{\boldsymbol{M}}(A')$ by Theorem V.

4.3. In what follows f_0 will always be assumed to be u.d. The discussion which follows could be based on an extension of the notion of $Tr_{\boldsymbol{M}}(A)$ (and of $Tr_{\boldsymbol{M}'}(A')$) to unbounded $A\eta\boldsymbol{M}$ (or $A'\eta\boldsymbol{M}'$) but we prefer an approach which avoids this.

THEOREM VII. *The sets* $Q_{f_0}(\boldsymbol{M})$ *and* $Q_{f_0}(\boldsymbol{M}')$ *are independent of the choice of of u.d.* f_0. *We will therefore denote them by* $Q(\boldsymbol{M})$ *and* $Q(\boldsymbol{M}')$, *respectively. Furthermore the values of* $\|Zf_0\|$, (Xf_0, Yf_0), *where* $X, Y, Z\epsilon Q(\boldsymbol{M})$, *or* $Q(\boldsymbol{M}')$ *are independent of the choice of the u.d.* f_0.

Owing to the symmetry between $\boldsymbol{M}$ and $\boldsymbol{M}'$ it suffices to prove the statements concerning $\boldsymbol{M}$; i.e., let f_0 and g_0 be u.d., then for $A\epsilon U(\boldsymbol{M})$, Af_0 is defined if and only if Ag_0 is defined and $\|Af_0\|=\|Ag_0\|$. Owing to the symmetry in f_0 and g_0 it will be sufficient to show that Ag_0 is defined if Af_0 is and $\|Af_0\|=\|Ag_0\|$.

By Lemma 4.2.2, there exists a unitary $U'\epsilon\boldsymbol{M}'$ such that $g_0=U'f_0$. Since A is $\epsilon U(\boldsymbol{M})$, $A=U'^{-1}AU'$. Hence since Af_0 exists $Af_0=U'^{-1}AU'f_0=U'^{-1}Ag_0$ exists. This implies that Ag_0 exists and since $Ag_0=U'Af_0$, $\|Ag_0\|=\|Af_0\|$. Also if $B\epsilon Q_{f_0}(\boldsymbol{M})$, then $Bg_0=U'Bf_0$ and $(Ag_0, Bg_0)=(U'Af_0, U'Bf_0)=(Af_0, Bf_0)$.

So we must have these mappings:

$$\text{(I)}\qquad \mathfrak{J}_{\boldsymbol{M}}:\mathfrak{H}\sim Q(\boldsymbol{M});\ \mathfrak{J}_{\boldsymbol{M}'}:\mathfrak{H}\sim Q(\boldsymbol{M}');\ \mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}:Q(\boldsymbol{M})\sim Q(\boldsymbol{M}').$$

By the corollary to Theorem V, $\mathfrak{J}_{\boldsymbol{M},\boldsymbol{M}'}$ maps $\boldsymbol{M}$ on $\boldsymbol{M}'$. So $\mathfrak{J}_{\boldsymbol{M}}$ maps $\boldsymbol{M}$ and $\mathfrak{J}_{\boldsymbol{M}'}$ maps $\boldsymbol{M}'$ on the same subset $\mathfrak{A}$ of $\mathfrak{H}$ and we have the further mappings

(II) $\mathfrak{J}_M: \mathfrak{A} \sim \boldsymbol{M}; \mathfrak{J}_{M'}: \mathfrak{A} \sim \boldsymbol{M}'; \mathfrak{J}_{M,M'}: \boldsymbol{M} \sim \boldsymbol{M}'.$

(II) is part of (I).

We now prove

LEMMA 4.3.1. *$\mathfrak{A}$ is a linear set, dense in $\mathfrak{H}$.*

Inasmuch as $\boldsymbol{M}$ is linear, $\mathfrak{A}$ is too.

Consider an $f \epsilon \mathfrak{H}$, $f = Zf_0$, $Z \epsilon Q(\boldsymbol{M})$. Now $Z = BU$, B self-adjoint and definite, $B \epsilon U(\boldsymbol{M})$, U unitary and $\epsilon \boldsymbol{M}$. Write B in its spectral form

$$B = \int_0^\infty \lambda dE(\lambda), \quad E(\lambda) \epsilon \boldsymbol{M}.$$

Then we have

$$[E(\mu)B] = \int_0^\mu \lambda dE(\lambda);$$

and therefore (i) $[E(\mu)B] \epsilon \boldsymbol{M}$ and (ii) if Bg is defined then we have strong $\lim_{\mu \to \infty} [E(\mu)B]g = Bg$. Thus $[E(\mu)B]U \epsilon \boldsymbol{M}$ and strong $\lim_{\mu \to \infty} [E(\mu)B]Uf_0 = BUf_0 = Zf_0 = f$. Since $[E(\mu)B]Uf_0 \epsilon \mathfrak{A}$, f is a condensation point of $\mathfrak{A}$. As this is true for any $f \epsilon \mathfrak{H}$, $\mathfrak{A}$ is dense in $\mathfrak{H}$.

DEFINITION 4.3.1. *If $f = Xf_0 = X'f_0$, $g = Yf_0 = Y'f_0$, X and $Y \epsilon Q(\boldsymbol{M})$, X' and $Y' \epsilon Q(\boldsymbol{M}')$, let $[[X]] = [[X']] = \|f\|$, $\langle\langle X, Y \rangle\rangle = \langle\langle X', Y' \rangle\rangle = (f, g)$.*

Theorem VII shows that the values of $[[X]]$, $[[X']]$, $\langle\langle X', Y' \rangle\rangle$, $\langle\langle X, Y \rangle\rangle$ are independent of the choice of the u.d. f_0.

LEMMA 4.3.2. *If X and Y are $\epsilon \boldsymbol{M}$, then $[[X]] = (Tr_M(X^*X))^{1/2} = (Tr_M(XX^*))^{1/2}$, $\langle\langle X, Y \rangle\rangle = Tr_M(Y^*X) = Tr_M(XY^*)$.*

As $[[X]] = (\langle\langle X, X \rangle\rangle)^{1/2}$, the second statement implies the first. Now clearly

$$\langle\langle X, Y \rangle\rangle = (Xf_0, Yf_0) = (Y^*Xf_0, f_0) = Tr_M(Y^*X)$$

since f is u.d. We also have $Tr_M(Y^*X) = Tr_M(XY^*)$ by §2.2, Property III, (vi). So $\langle\langle X, Y \rangle\rangle = Tr_M(Y^*X) = Tr_M(XY^*)$.

LEMMA 4.3.3. *If X', $Y' \epsilon \boldsymbol{M}'$, then $[[X']] = (Tr_{M'}(X'^*X'))^{1/2} = (Tr_{M'}(X'X'^*))^{1/2}$, $\langle\langle X', Y' \rangle\rangle = Tr_{M'}(Y'^*X') = Tr_M(X'Y'^*)$.*

Replace $\boldsymbol{M}$ by $\boldsymbol{M}'$ in the proof of Lemma 4.3.2.

THEOREM VIII. *If we use the definitions*

$$\langle\langle X, Y \rangle\rangle = Tr_M(Y^*X) = Tr_M(XY^*),$$
$$[[X]] = (\langle\langle X, X \rangle\rangle)^{1/2} = (Tr_M(X^*X))^{1/2} = (Tr_M(XX^*))^{1/2},$$

then $\boldsymbol{M}$ is an incomplete Hilbert space. Its completion in the usual (Cantor) way gives a Hilbert space $\tilde{\boldsymbol{M}}$ which may be identified with $Q(\boldsymbol{M})$.

This is clear by the isomorphisms (I) and (II) and Lemmas 4.3.1 and 4.3.2. $\boldsymbol{M}$ and $\tilde{\boldsymbol{M}}$, that is, $Q(\boldsymbol{M})$, are isomorphic to $\mathfrak{A}$ and $\mathfrak{H}$ respectively, using the isomorphism $\mathfrak{I}_{\boldsymbol{M}}$.

The corresponding facts hold for $\boldsymbol{M}'$; their formulation is obvious.

THEOREM IX. *$X \epsilon Q(\boldsymbol{M})$ implies $X^* \epsilon Q(\boldsymbol{M})$ and $[[X]] = [[X^*]]$.*

Assume $X \epsilon Q(\boldsymbol{M})$, i.e., $X \epsilon U(\boldsymbol{M})$; Xf_0 is defined. Write $X = BU$, B self-adjoint and $\eta \boldsymbol{M}$, U unitary and $\epsilon \boldsymbol{M}$. Thus BUf_0 is defined. Now Uf_0 is u.d. (Lemma 4.2.1), hence the existence of BUf_0 implies $B \epsilon Q(\boldsymbol{M})$ (Theorem VII). This in turn implies the existence of Bf_0 and with it $X^* f_0 = U^{-1} B f_0$. So $X^* \epsilon Q(\boldsymbol{M})$.

As f_0 and Uf_0 are both u.d. (Lemma 4.2.1) we have by Definition 4.3.1, $[[X]] = \|Xf_0\| = \|BUf_0\| = [[B]] = \|Bf_0\| = \|U^{-1}Bf_0\| = \|X^*f_0\| = [[X^*]]$.

COROLLARY. *$X \sim X^*$ is an involutory conjugate anti-isomorphism of $Q(\boldsymbol{M})$ and isometric.*

$X \sim X^*$ maps $Q(\boldsymbol{M})$ on part of itself by Theorem IX. This and $X^{**} = X$ show that $X \sim X^*$ is a one-to-one and involutory mapping of $Q(M)$ on itself. It is isometric by Theorem IX; $[[X]] = [[X^*]]$ and it is obviously a conjugate anti-isomorphism.

The corresponding facts hold for $\boldsymbol{M}'$; the formulation is obvious.

4.4. We next discuss the algebra of $Q(\boldsymbol{M})$.

PROPERTY I^0. *A, $B \epsilon Q(\boldsymbol{M})$ imply αA, A^*, $[A+B] \epsilon Q(\boldsymbol{M})$. If also either A or B is $\epsilon \boldsymbol{M}$, then $[AB]$ is $\epsilon Q(\boldsymbol{M})$.*

αA, $[A+B]$ are $\epsilon Q(\boldsymbol{M})$ since $Q(\boldsymbol{M})$ is linear (along with $\mathfrak{H}$). A^* is $\epsilon Q(\boldsymbol{M})$ by Theorem IX.

Assume now that $A \epsilon \boldsymbol{M}$, $B \epsilon Q(\boldsymbol{M})$. Then Bf_0 is defined and so is $ABf_0 = [AB]f_0$. Thus $[AB] \epsilon Q(\boldsymbol{M})$. If $A \epsilon Q(\boldsymbol{M})$, $B \epsilon \boldsymbol{M}$, then $B^* \epsilon \boldsymbol{M}$, $A^* \epsilon Q(\boldsymbol{M})$, so $[B^*A^*] \epsilon Q(\boldsymbol{M})$. Since $[B^*A^*]^* = [AB]$ (R.O., Theorem XV), $[AB]$ is $\epsilon Q(\boldsymbol{M})$.

PROPERTY II^0. (i) $[[\alpha A]] = |\alpha| \cdot [[A]]$, (ii) $[[A^*]] = [[A]]$, (iii) $[[A+B]] \leqq [[A]] + [[B]]$, (iv) $[[[AB]]] \leqq |||A||| \cdot [[B]]$ *and* $[[A]] \cdot |||B|||$.

(i) and (iii) hold because $Q(\boldsymbol{M})$ is isomorphic to $\mathfrak{H}$. (ii) holds by Theorem IX.

Consider (iv). We have

$$[[[AB]]] = \|[AB]f_0\| = \|ABf_0\| \leqq |||A||| \cdot \|Bf_0\| = |||A||| \cdot [[B]],$$

proving the first inequality. The second follows from this, by using (ii)

$$[[[AB]]] = [[[AB]^*]] = [[[B^*A^*]]] \leqq |||\, B^* \,||| \cdot [[A^*]] = |||\, B \,||| \cdot [[A]].$$

We have $\mathfrak{H} \sim Q(\boldsymbol{M})$. Neither $\mathfrak{H}$ nor $Q(\boldsymbol{M})$ depends on the choice of the u.d. f_0 but $\mathfrak{J}_M$ does. This dependence is as follows.

Property III^0. *Let us replace the u.d. f_0 by a u.d. g_0 and let $\mathfrak{F}_M$ denote the resulting correspondence. Let U be unitary and $\epsilon \boldsymbol{M}$ and such that $Uf_0 = g_0$* (cf. Lemma 4.2.1). *Then if $X = YU$, X and $Y \epsilon Q(\boldsymbol{M})$, we have*

$$\mathfrak{J}_M : f \sim X; \qquad \mathfrak{F}_M : f \sim Y$$

if $f = Xf_0$.

This is clear since $f = Xf_0 = YUf_0 = Yg_0$.

We can now determine another notion which does not depend on the choice of the u.d. f_0.

Property IV^0. *The linear set $\mathfrak{A}$* (cf. §4.3, the correspondences (II)) *does not depend on the choice of the u.d. f_0.*

This follows from the definition of $\mathfrak{A}$ and the fact that if $X = YU$, U unitary and either X or Y is bounded, then both X and Y are bounded.

Now $Q(\boldsymbol{M}) \sim Q(\boldsymbol{M}')$ under $\mathfrak{J}_{M,M'}$ and while neither $Q(\boldsymbol{M})$ nor $Q(\boldsymbol{M}')$ depends on the choice of the u.d. f_0 yet $\mathfrak{J}_{M,M'}$ does. We now obtain this dependence.

Property V^0. *Let the u.d. f_0 be replaced by the u.d. g_0 and let the resulting correspondence between $Q(\boldsymbol{M})$ and $Q(\boldsymbol{M}')$ be denoted by $\mathfrak{F}_{M,M'}$. Then if $U \epsilon \boldsymbol{M}$ is unitary and such that $g_0 = Uf_0$, and $X = U^{-1}YU$, X and $Y \epsilon Q(\boldsymbol{M})$, then*

$$\mathfrak{J}_{M,M'} : X' \sim X; \qquad \mathfrak{F}_{M,M'} : X' \sim Y$$

if $X'f_0 = Xf_0$, $X' \epsilon Q(\boldsymbol{M}')$.

Let $X'f_0 = Xf_0$, then $X'g_0 = X'Uf_0 = UX'f_0 = UXf_0 = UXU^{-1}g_0$.

Definition 4.4.1. *The isomorphism $\mathfrak{H} \sim Q(\boldsymbol{M})$ makes correspond to every operator P in $\mathfrak{H}$ an operator P° in $Q(\boldsymbol{M})$.*

Observe that inasmuch as the elements of $Q(\boldsymbol{M})$ are operators in $\mathfrak{H}$, P° is an operator on the operators of $\mathfrak{H}$.

Theorem X. (i) *If $A \epsilon \boldsymbol{M}$, then*

$$A^\circ Z = [AZ];$$

(ii) *if $A' \epsilon \boldsymbol{M}'$ and $A' \sim A \epsilon \boldsymbol{M}$ under $\mathfrak{J}_{M,M'}$, then*

$$A'^\circ = [ZA].$$

P° is defined as follows. If $g = Xf_0$, $h = Pg$, $h = Yf_0$, X and $Y \epsilon Q(\boldsymbol{M})$, then $P^\circ X = Y$. Thus to prove (i), we have $(A^\circ Z)f_0 = A(Zf_0) = [AZ]f_0$ and to show (ii) $(A'^\circ Z)f_0 = A'Zf_0 = ZA'f_0 = ZAf_0 = [ZA]f_0$.

COROLLARY 1. *Let $\boldsymbol{M}^0$ be the set of all operators L_A in $Q(\boldsymbol{M})$, $A \epsilon \boldsymbol{M}$, where $L_A Z = [AZ]$, and $\boldsymbol{M}_0$ the set of all operators R_A in $Q(\boldsymbol{M})$, $A \epsilon \boldsymbol{M}$, where $R_A Z = [ZA]$. Then $\mathfrak{J}_M$ carries $\boldsymbol{M}$ into $\boldsymbol{M}^0$ and $\boldsymbol{M}'$ into $\boldsymbol{M}_0$.*

This is obvious by Theorem X.

COROLLARY 2. *$\boldsymbol{M}^0$, $\boldsymbol{M}_0$ are rings and $\boldsymbol{M}^{0\prime} = \boldsymbol{M}_0$ (all in $Q(\boldsymbol{M})$). They are factors of class* $(\mathrm{II}_1, \mathrm{II}_1)$ *with* $C = 1$.

Since $\mathfrak{J}_M$ is a spatial isomorphism of $\mathfrak{H}$ and $Q(\boldsymbol{M})$ which takes $\boldsymbol{M}$, $\boldsymbol{M}'$ into $\boldsymbol{M}^0$, $\boldsymbol{M}_0$ respectively, these properties hold.

We have now characterized $\mathfrak{H}$, $\boldsymbol{M}$, $\boldsymbol{M}'$ by the situation in $Q(\boldsymbol{M})$, $\boldsymbol{M}^0$, $\boldsymbol{M}_0$ to which they are spatially isomorphic. The spatial isomorphism is $\mathfrak{J}_M$ which depends on an arbitrary u.d. f_0, while $Q(\boldsymbol{M})$, $\boldsymbol{M}^0$, $\boldsymbol{M}_0$ themselves do not. All the influence of the choice of f_0 is however a further transformation R_U, $X = YU$ ($U \epsilon \boldsymbol{M}$, U unitary so $R_U \epsilon \boldsymbol{M}_0$). $\mathfrak{J}_M$ always carries the totality of u.d. elements g_0 of $\mathfrak{H}$ into the totality of unitary elements $U \epsilon Q(\boldsymbol{M})$, f_0 being that element which corresponds to 1.

4.5. The following important isomorphism theorem can now be proved.

THEOREM XI. *Let $\mathfrak{H}_1$ and $\mathfrak{H}_2$ be two Hilbert spaces and $\boldsymbol{M}_1$, $\boldsymbol{M}_1'$ and $\boldsymbol{M}_2$, $\boldsymbol{M}_2'$ respectively be factor pairs in them, both of class* $(\mathrm{II}_1, \mathrm{II}_1)$ *and with $C = 1$ in the standard normalization. Then an algebraic ring isomorphism of $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ is a necessary and sufficient condition for the existence of a spatial isomorphism of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ which takes $\boldsymbol{M}_1$, $\boldsymbol{M}_1'$ into $\boldsymbol{M}_2$, $\boldsymbol{M}_2'$.*

The necessity is obvious and so we prove the sufficiency. Let $\mathfrak{F}$ be the algebraic ring isomorphism of $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$. By §2.2, Property IV, $Tr_{M_1}(A_1) = Tr_{M_2}(A_2)$ if $A_1 \sim A_2$ under $\mathfrak{F}$. So $[[A_1]] = [[A_2]]$, $\langle\langle A_1, B_1\rangle\rangle = \langle\langle A_2, B_2\rangle\rangle$, if $A_1 \sim A_2$, $B_2 \sim B_2$ under $\mathfrak{F}$. Thus $\mathfrak{F}$ is an isometric mapping of $\boldsymbol{M}_1$ on $\boldsymbol{M}_2$ and therefore extends by continuity in a unique way to a linear and isometric mapping of $Q(\boldsymbol{M})$ on $Q(\boldsymbol{M}')$ which we call $\mathfrak{F}$ again. $A_1 \sim A_2$, $B_1 \sim B_2$ under $\mathfrak{F}$ imply $A_1B_1 \sim A_2B_2$ under $\mathfrak{F}$ if A_1, $B_1 \epsilon \boldsymbol{M}_1$ (and thus A_2, $B_2 \epsilon \boldsymbol{M}_2$). By continuity this will even hold, if one of A_1, B_1 is in $\boldsymbol{M}_1$ and the other merely in $Q(\boldsymbol{M}_1)$ (and one of A_2, B_2 in $\boldsymbol{M}_2$, and the other merely in $Q(\boldsymbol{M}_2)$). So $\mathfrak{F}$ is a spatial isomorphism of $Q(\boldsymbol{M}_1)$ and $Q(\boldsymbol{M}_2)$ which carries $\boldsymbol{M}_1^0$, $\boldsymbol{M}_{1,0}$ into $\boldsymbol{M}_2^0$, $\boldsymbol{M}_{2,0}$ (by Corollary 1 to Theorem XI). Now (by the same corollary) $\mathfrak{J}_{M_2}\mathfrak{F}\mathfrak{J}_{M_1}^{-1}$ is a spatial isomorphism of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ which carries $\boldsymbol{M}_1$, $\boldsymbol{M}_1'$ into $\boldsymbol{M}_2$, $\boldsymbol{M}_2'$.

Theorem XI could be extended to cover cases $(\mathrm{II}_1, \mathrm{II}_1)$, where $C \gtreqless 1$ in the standard normalization as well as other combinations of II_1 and II_∞. In all

cases a reduction of the spatial-isomorphism questions of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ to algebraic-isomorphism questions of $\mathfrak{H}_1$ and $\mathfrak{H}_2$ (plus the behavior of C) results. We will discuss these questions in detail in later publications.

Appendix

1. Consider a ring $\boldsymbol{M}$ of class II_1 in $\mathfrak{H}$. Then $\boldsymbol{M}'$ is of class II_1 or II_∞. Normalize $D_{\boldsymbol{M}}$ and $D_{\boldsymbol{M}'}$ so as to have $D_{\boldsymbol{M}}(\mathfrak{H})=1$ and $C=1$, so $D_{\boldsymbol{M}'}(\mathfrak{H})=\alpha$ (cf. §1.1). By choosing an $n=1, 2, \cdots$ with $1/n \leqq \alpha$ and then an $\mathfrak{M}'\eta \boldsymbol{M}'$ with $D_{\boldsymbol{M}'}(\mathfrak{M}')=1/n$, apply R.O., §11.3, to form $\boldsymbol{M}_{(\mathfrak{M}')}$, $\boldsymbol{M}'_{(\mathfrak{M}')}$ in $\mathfrak{M}'$. Then $\boldsymbol{M}_{(\mathfrak{M}')}$ is algebraically-ring-isomorphic to $\boldsymbol{M}$, and $D_{\boldsymbol{M}(\mathfrak{M}')}(\mathfrak{M}')=D_{\boldsymbol{M}}(\mathfrak{H})=1$ (cf. R.O., Lemmas 11.3.3 and 11.3.6). So if we are interested in the algebraical properties of $\boldsymbol{M}$ only, we may assume without any loss in generality that $\boldsymbol{M}$, $\boldsymbol{M}'$ are in case II_1, II_1 and that $\alpha=1/n$, $n=1, 2, \cdots$. Or, in standard normalization $C=n$, $n=1, 2, \cdots$ (cf. R.O., Theorem X).

Now form the direct product of $\mathfrak{H}$ with an n-dimensional Euclidean space $E_n \oplus \mathfrak{H}$ as in §2.1, for the case $C>1$, and consider $\boldsymbol{M}^{(2)}$ in $E_n \oplus \mathfrak{H}$. The argument used in §2.1 shows that $\boldsymbol{M}^{(2)\prime}=R(\boldsymbol{N}^{(1)}, \boldsymbol{M}^{(2)\prime})$ and $C^{(2)}=C/n=1$ and $\boldsymbol{M}^{(2)}$ is ring isomorphic to $\boldsymbol{M}$.

So we have for $\boldsymbol{M}^{(2)}$, $\boldsymbol{M}^{(2)\prime}$ in the standard normalization $C=1$. If we are therefore interested in the algebraical properties of $\boldsymbol{M}$ only, we may even assume without any loss in generality that $C=1$ in standard normalization; that is, $\alpha=1$ in the normalization of §1.1. We will assume this in what follows.

It may be noticed concerning subrings that the metric of $Q(\boldsymbol{M})$ (cf. Definition 4.3.1) is determined by the algebra of $\boldsymbol{M}$ (cf. §2.2). A subring demands closure in the weak operator topology, but it will be shown elsewhere that for subrings of $\boldsymbol{M}$ weak, relative closure in $\boldsymbol{M}$ for the $[[X-Y]]$ metric is equivalent to closure in the weak topology.

2. Under these conditions there is an analogy between $\boldsymbol{M}$ and the matrices of a Euclidean space E_n, described by an interesting Lebesgue-Stieltjes-Radon measure in the plane.

We can use the results of §§4.2–4.4, and thus we may form the Hilbert-space $Q(\boldsymbol{M})$ which is isomorphic to $\mathfrak{H}$, and in which $\boldsymbol{M}$, $\boldsymbol{M}'$ are located by Theorem X.

Let $E(\lambda)$, $a<\lambda<b$, be a resolution of unity, all $E(\lambda)\epsilon\boldsymbol{M}$. Put $A=\int_{-\infty}^{\infty}\lambda dE(\lambda)$ (in the well known symbolic sense). For any Borel-set of real numbers S form

$$e_S(\lambda) = \begin{cases} 1 \text{ for } \lambda\epsilon S \\ 0 \text{ for } \lambda \not\epsilon S, \end{cases}$$

and

$$E(S) = e_S(A) = \int_{-\infty}^{\infty} e_S(\lambda)dE(\lambda) = \int_S dE(\lambda)$$

(symbolically). As $\overline{e_S(\lambda)}=e_S(\lambda)=(e_S(\lambda))^2$, so $e_S(A)=e_S(A)^*=(e_S(A))^2$; that is, $E(S)=e_S(A)\epsilon \boldsymbol{M}$ is a projection. (Cf. also Maeda, Journal of Science, Hiroshima University, Ser. A, vol. 4 (1934), pp. 57–91.)

Consider now point sets in the λ, μ-plane P and in particular sets of the form $S_1\otimes S_2$:

$$(\lambda, \mu)\epsilon S_1 \otimes S_2 \quad \text{means} \quad \lambda\epsilon S_1, \ \mu\epsilon S_2,$$

S_1, S_2 being two Borel-sets of real numbers. Define for any $X\epsilon \boldsymbol{M}$ (or even $X\epsilon Q(\boldsymbol{M})$),

$$\begin{aligned} X_{S_1\times S_2} &= E(S_1)XE(S_2) \\ \omega(X;\, S_1 \otimes S_2) &= [[X_{S_1\times S_2}]]^2 = Tr((X_{S_1\times S_2})^* X_{S_1\times S_2}) \\ &= Tr(E(S_2)X\ \ E(S_1)\cdot E(S_1)XE(S_2)) \\ &= Tr(E(S_2)X^*\cdot E(S_1)XE(S_2)) \\ &= Tr(E(S_1)XE(S_2)\cdot E(S_2)X^*) \\ &= Tr(E(S_1)XE(S_2)X^*). \end{aligned}$$

The first expression for $\omega(X, S_1\otimes S_2)$ shows, that it is always $\geqq 0$, the last one, that it is a totally additive set-function of S_1 and of S_2.

Denote the set of all λ by Λ, then the above facts imply:

$$\begin{aligned} \omega(X; S_1 \otimes S_2) &\leqq \omega(X; S_1 \otimes S_2) + \omega(X; (\Lambda - S_1) \otimes S_2) \\ &= \omega(X; \Lambda \otimes S_2) \\ &\leqq \omega(X; \Lambda \otimes S_2) + \omega(X; \Lambda \otimes (\Lambda - S_2)) \\ &= \omega(X; \Lambda \otimes \Lambda) = [[X_{\Lambda\times\Lambda}]]^2 \\ &= [[1\cdot X\cdot 1]]^2 = [[X]]^2. \end{aligned}$$

So we have

LEMMA A. (i) *$\omega(X; S_1\otimes S_2)$ is defined for all linear Borel-sets S_1, S_2.* (ii) *It is totally additive in S_1 as well as in S_2.* (iii) *We have always*

$$0 \leqq \omega(X; S_1 \otimes S_2) \leqq [[X]]^2 = Tr(X^*X).$$

3. We can use Lemma A to define a Lebesgue-Stieltjes-Radon measure $\mu(X; T)$ for all plane Borel-sets $T(\subset P)$ with the help of $\omega(X; S_1\otimes S_2)$.

DEFINITION A. *If T is a plane Borel-set ($T \subset P$), then consider all sequences of linear Borel-set $S_1^{(i)}$, $S_2^{(i)}$, $i=1, 2, \cdots$ (all $S_1^{(i)}$, $S_2^{(i)} \subset\Lambda$) for which*

$$(*) \qquad T \subset \sum_{i=1}^{\infty} S_1^{(i)} \otimes S_2^{(i)}.$$

For every such sequence form the (numerical) sum

$$(**) \qquad \sum_{i=1}^{\infty} \omega(X; S_1^{(i)} \otimes S_2^{(i)}).$$

Denote the g.l.b. of all numbers $(**)$ *by* $\mu(X; T)$.

We prove now

LEMMA B. (i) $\mu(X; T)$ *is defined for all plane Borel-sets* $T \subset P$.

(ii) *It is totally additive in* T.

(iii) *We have always*

$$0 \leqq \mu(X; T) \leqq [[X]]^2 = Tr(X^*X).$$

(iv) *In particular*

$$\mu(X; S_1 \otimes S_2) = \omega(X; S_1 \otimes S_2)$$

and especially

$$\mu(X, P) = [[X]]^2 = Tr(X^*X).$$

(i) is obvious. To prove (ii), observe first that our Definition A makes

$$\mu(X; T_1 + T_2 + \cdots) \leqq \mu(X; T_1) + \mu(X; T_2) + \cdots$$

obvious, so we need to prove

$$\mu(X; T_1 + T_2 + \cdots) \geqq \mu(X; T_1) + \mu(X; T_2) + \cdots$$

only when $T_i \cdot T_j = 0$ for $i \neq j$.

Even $\mu(X; T_1+T_2) \geqq \mu(X; T_1)+\mu(X; T_2)$ suffices. Then finite induction gives $\mu(X; T_1+T_2+ \cdots +T_n) \geqq \mu(X; T_1)+\mu(X; T_2)+ \cdots +\mu(X; T_n)$, and so $\mu(X; T_1+T_2+ \cdots) \geqq \mu(X; T_1+ \cdots +T_n) \geqq \mu(X; T_1)+ \cdots +\mu(X; T_n)$ and as this holds for all $n=1, 2, \cdots$, it implies $\mu(X; T_1+T_2+ \cdots) \geqq \mu(X; T_1)+\mu(X; T_2)+ \cdots$.

We will prove

$$(\S) \qquad \mu(X; T) = \mu(X; TU) + \mu(X; T - TU)$$

for all T, U this gives our above inequality if $T=T_1+T_2$, $U=T_1$. Call a U, for which (§) holds for all plane Borel-sets T, following Carathéodory (cf. (4), pp. 246–252), measurable. We must show that all U are measurable.

If $U=S_1 \otimes S_2$, then $T \subset \sum_{i=1}^{\infty} S_1^{(i)} \otimes S_2^{(i)}$ implies $TU \subset \sum_{i=1}^{\infty} S_1^{(i)} S_1 \otimes S_2^{(i)} S_2$, $T-TU \subseteq \sum_{i=1}^{\infty} S_1^{(i)} S_1 \otimes (S_2^{(i)} - S_2^{(i)} S_2) + \sum_{i=1}^{\infty} (S_1^{(i)} - S_1^{(i)} S_1) \otimes S_2^{(i)}$. Thus our Definition A gives immediately (§) with $\geqq$ in it, and as $\leqq$ is obvious (cf. above) it proves (§). So all $U=S_1 \otimes S_2$ are measurable, and therefore in particular all plane intervals (=rectangles).

Now the measurable sets U form a Borel-ring (cf., for instance, (4), loc. cit.). These considerations apply literally to the present case. Thus every plane Borel-set U is measurable. This completes the proof.

We now prove (iii). $\mu(X;\,T) \geqq 0$ because all expressions (* *) in Definition A are $\geqq 0$. The relation $\mu(X;\,T) \leqq [[X]]^2 = Tr(X^*X)$ results by putting $S_1^{(1)} = S_2^{(1)} = \Lambda$ and all other $S_1^{(i)} = S_2^{(i)} = 0$.

To prove (iv), put $S_1^{(1)} = S_1$, $S_2^{(1)} = S_2$ and all other $S_1^{(i)} = S_2^{(i)} = 0$. This gives $\mu(X;\,S_1 \otimes S_2) \leqq \omega(X;\,S_1 \otimes S_2)$. So we must prove $\geqq$ only; that is,

$$S_1 \otimes S_2 \subset \sum_{i=1}^{\infty} S_1^{(i)} \otimes S_2^{(i)} \quad \text{implies} \quad \omega(X;S_1 \otimes S_2) \leqq \sum_{i=1}^{\infty} \omega(X;S_1^{(i)} \otimes S_2^{(i)}).$$

Considering the properties of $\omega(X;\,S_1 \otimes S_2)$ given in Lemma A, this implication follows literally as in the paper of Łomnicki and Ulam (Fundamenta Mathematicae, vol. 23 (1934), pp. 237–278; cf. also the lecture notes of the second-named author for the year 1934–1935).

So we have proved $\mu(X;S_1 \otimes S_2) = \omega(X;S_1 \otimes S_2)$. Put now $S_1 = S_2 = \Lambda$; then

$$\mu(X;\,P) = \mu(X;\Lambda \otimes \Lambda) = \omega(X;\Lambda \otimes \Lambda) = [[X]]^2 = Tr(X^*X).$$

results.

4. The plane measure $\mu(X;\,T)$ may be used to "locate" the "position" of X in the λ, μ-plane P. It gives the entire plane a total "measure" or "weight" $[[X]]^2$ (which is >0 if $X \neq 0$), and any part T of it correspondingly a $\mu(X;\,T)$ ($\geqq 0$). It plays the same role, as the sum of the absolute-value-squares of all matrix-elements in a certain area T in the matrix-scheme (which may be looked at as a region in the plane) for the matrices of a finite-dimensional Euclidean space E_n (that is, $\boldsymbol{M}$ in a case (I_n), $n = 1, 2, \cdots$).

We will analyse it more thoroughly in subsequent publications.

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

GENERALIZED WEIGHT PROPERTIES OF THE RESULTANT OF $n+1$ POLYNOMIALS IN n INDETERMINATES†

BY
OSCAR ZARISKI

1. **Introduction.** *The multiplicity of intersection of two plane algebraic curves, $f(x, y)=0$ and $g(x, y)=0$, at a common point $O(a, b)$, r-fold for f and s-fold for g, is not less than rs, and is greater than rs if and only if the two curves have in common a principal tangent at O.* The standard proof of this well known theorem of the theory of higher plane curves makes use of Puiseux expansions. If, namely, $R(x)=R(f, g)$ denotes the resultant of f and g, considered as polynomials in y, and if $y_1, y_2, \cdots, y_n$ and $\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_m$ are the roots of $f=0$ and $g=0$ respectively, then, the axes being in generic position, the intersection multiplicity at O is defined as the multiplicity of the root $x=a$ of the resultant $R(x)$, and this multiplicity is found by substituting into the product $\prod_{i=1}^{n}\prod_{j=1}^{m}(y_i-\bar{y}_j)$ the Puiseux expansions of the roots y_i and $\bar{y}_j$. A less known proof, in which the multiplicity to which the factor $x-a$ occurs in $R(x)$ is derived in a purely algebraic manner, was given by C. Segre.‡ Following a procedure due to A. Voss,§ Segre uses the Sylvester determinant and arrives at the required result by a skillful manipulation of the rows and columns.

In the first part of this paper (§§2, 3), we give a new proof of the property of the resultant $R(f, g)$ (see Theorem 1), which is implicitly contained in the quoted paper by C. Segre and of which the above intersection theorem is an immediate corollary. This proof makes use only of the intrinsic properties of the resultant and so contains the germ of an extension to the case of $n+1$ polynomials in n variables. In the second part (§§4–9) we extend Theorem 1 to the resultant of $n+1$ polynomials (Theorem 6). From Theorem 6 follows as a corollary the analogous intersection theorem for hypersurfaces in S_{n+1} (§9).

I. Two polynomials in one variable

2. **A generalized weight property of the resultant.** Let

---

† Presented to the Society, December 31, 1936; received by the editors June 18, 1936.

‡ C. Segre, *Le molteplicità nelle intersezioni delle curve piane algebriche con alcune applicazioni ai principi della teoria di tali curve*, Giornale de Matematiche di Battaglini, vol. 36 (1898).

§ A. Voss, *Über einen Fundamentalsatz aus der Theorie der algebraischen Functionen*, Mathematische Annalen, vol. 27 (1886).

$$f = a_0y^n + a_1y^{n-1} + \cdots + a_n,$$
$$g = b_0y^m + b_1y^{m-1} + \cdots + b_m,$$

be two polynomials, with literal coefficients, and let $R(f, g)$ be their resultant:

$$R(f, g) = \sum a_0^{i_0} a_1^{i_1} \cdots a_n^{i_n} b_0^{j_0} b_1^{j_1} \cdots b_m^{j_m},$$

where, by well known properties of R, we have

$$i_0 + i_1 + \cdots + i_n = m, \quad j_0 + j_1 + \cdots + j_m = n,$$
$$i_1 + 2i_2 + \cdots + ni_n + j_1 + 2j_2 + \cdots + mj_m = mn.$$

THEOREM 1. *Let r and s be two non-negative integers, $r \leqq n$, $s \leqq m$. If we give to each coefficient a_i (b_j) the weight $r-i$ ($s-j$) or zero, according as $r-i \geqq 0$ ($s-j \geqq 0$) or $r-i \leqq 0$ ($s-j \leqq 0$), then the weight of any term in the resultant $R(f, g)$ is $\geqq rs$. The sum of terms of weight rs is given by the following expression*:

$$(-1)^{(m-s)r} R(f_r, g_s) R(f^*_{n-r}, g^*_{m-s}),$$

where

$$f_r = a_0y^r + \cdots + a_r, \qquad f^*_{n-r} = a_ry^{n-r} + \cdots + a_n;$$
$$g_s = b_0y^s + \cdots + b_s, \qquad g^*_{m-s} = b_sy^{m-s} + \cdots + b_m.$$

We consider the polynomials

$$\bar{f} = a_0t^ry^n + \cdots + a_{r-1}ty^{n-r+1} + a_ry^{n-r} + \cdots + a_n,$$
$$\bar{g} = b_0t^sy^m + \cdots + b_{s-1}ty^{m-s+1} + b_sy^{m-s} + \cdots + b_m,$$

where t is a new indeterminate. Let t^k be the highest power of t which divides the resultant $R(\bar{f}, \bar{g})$, $\bar{f}$ and $\bar{g}$ being considered as polynomials in y:

$$R(\bar{f}, \bar{g}) = t^kR_1(a_i, b_j, t), \qquad R_1(a_i, b_j, 0) \neq 0. \tag{1}$$

By a well known property of the resultant, we have $R(\bar{f}, \bar{g}) = A\bar{f} + B\bar{g}$, where A and B are polynomials in y, a_i, b_j, t, with integral coefficients; or using a familiar notation: $R(\bar{f}, \bar{g}) \equiv 0(\bar{f}, \bar{g})$. We put $\bar{f} = \bar{f}^* + a_n$, $\bar{g} = \bar{g}^* + b_m$. If we make the substitution $a_n = -\bar{f}^*$, $b_m = -\bar{g}^*$ in the identity $R(\bar{f}, \bar{g}) = A\bar{f} + B\bar{g}$, then $\bar{f}$ and $\bar{g}$ vanish, and therefore also $t^kR_1(a_i, b_j, t)$ must vanish. Since t is unaltered by the substitution, we have

$$R_1(a_0, \cdots, a_{n-1}, -\bar{f}^*; b_0, \cdots, b_{m-1}, -\bar{g}^*; t) = 0.$$

If we now order $R_1(a_i, b_j, t)$ according to the powers of $a_n + \bar{f}^*$ and $b_m + \bar{g}^*$, the constant term vanishes, and hence $R_1(a_i, b_j, t) \equiv 0(\bar{f}, \bar{g})$.† Putting $t = 0$,

† This proof that $t^kR_1 \equiv 0(\bar{f}, \bar{g})$ implies $R_1 \equiv 0(\bar{f}, \bar{g})$ is taken from van der Waerden, *Moderne Algebra*, II, p. 15 (quoted in the sequel as W.) Further on we shall use frequently the notion and properties of *inertia forms* as given in W., pp. 15–21.

we find $R_{10}=R_1(a_i, b_j, 0)\equiv 0(f^*_{n-r}, g^*_{m-s})$, and hence R_{10} vanishes whenever f^*_{n-r} and g^*_{m-s} have a common factor of degree $\geqq 1$ in y. Consequently, by a well known property of the resultant, *R_{10} is divisible by $R(f^*_{n-r}, g^*_{m-s})$*, provided, however, that $n\neq r$ and $m\neq s$ (inequalities assuring the irreducibility of $R(f^*_{n-r}, g^*_{m-s})$).

If we now consider the resultant $R(\bar{f}, \bar{g})$ of the following polynomials:

$$\bar{f} = a_0y^n + \cdots + a_ry^{n-r} + ta_{r+1}y^{n-r-1} + \cdots + t^{n-r}a_n,$$
$$\bar{g} = b_0y^m + \cdots + b_sy^{m-s} + tb_{s+1}y^{m-s-1} + \cdots + t^{m-s}b_m,$$

or, what is the same, the resultant of the polynomials

$$a_nt^{n-r}y^n + \cdots + a_{r+1}ty^{r+1} + a_ry^r + \cdots + a_0,$$
$$b_mt^{m-s}y^m + \cdots + b_{s+1}ty^{s+1} + b_sy^s + \cdots + b_0,$$

and if we put $R(\bar{f}, \bar{g}) = t^lR_2(a_i, b_j, t)$ and $R_{20}=R_2(a_i, b_j, 0)$, where t^l is the highest power of t which divides $R(\bar{f}, \bar{g})$, we conclude as before that R_{20} is divisible by the resultant of the polynomials

$$a_ry^r + a_{r-1}y^{r-1} + \cdots + a_0,$$
$$b_sy^s + b_{s-1}y^{s-1} + \cdots + b_0;$$

i.e., R_{20} is divisible by $R(f_r, g_s)$, provided $r\neq 0$ and $s\neq 0$. But since

$$\bar{f} = t^{n-r}\bar{f}\left(\frac{y}{t}\right), \qquad \bar{g} = t^{m-s}\bar{g}\left(\frac{y}{t}\right),$$

we have

$$t^{rs}R(\bar{f}, \bar{g}) = t^{(n-r)(m-s)}R(\bar{f}, \bar{g}),$$

i.e., $R(\bar{f}, \bar{g})$ and $R(\bar{f}, \bar{g})$ differ only by a factor which is a power of t. Hence $R_{20}=R_{10}$, and therefore *R_{10} is divisible by both $R(f^*_{n-r}, g^*_{m-s})$ and $R(f_r, g_s)$.*

Assuming that $r\neq 0, n, s\neq 0, m$, we have that $R(f^*_{n-r}, g^*_{m-s})$ and $R(f_r, g_s)$ are irreducible and distinct polynomials in the coefficients a_i, b_j. [a_0, for instance, actually occurs in $R(f_r, g_s)$, but does not occur in $R(f^*_{n-r}, g_{m-s})$.] Hence *R_{10} is divisible by the product $R(f_r, g_s)\cdot R(f^*_{n-r}, g^*_{m-s})$.* Since $R(\bar{f}, \bar{g})$, and hence also R_{10}, is of degree m in the coefficients of f and of degree n in the coefficients of g, we conclude that

$$R_{10} = cR(f_r, g_s)R(f^*_{n-r}, g^*_{m-s}),$$

where c is a numerical factor (an integer).

Assume $r=0, s\neq m$. Then $f^*_{n-r}=f$ and $R(f^*_{n-r}, g^*_{m-s})$ is irreducible and of degree n in the coefficients of g, and consequently the quotient $R_{10}/R(f^*_{n-r}, g^*_{m-s})$ is independent of the coefficients of g. On the other hand, $R(\bar{f}, \bar{g})$ contains

the term $a_0^m b_m^n$, so that the exponent k of t in (1) equals 0, and therefore R_{10} can vanish only if f and g^*_{m-s} have a common zero or if $a_0=0$. It follows that also in this case $R_{10}=ca_0^s R(f, g^*_{m-s})=c\cdot R(f_0, g_s)\, R(f, g^*_{m-s})$, where c is an integer.

The case $s=0$, $r\neq n$ is treated in a similar manner.

R_{10} is visibly given by the product $a_0^m b_m^n=R(f_0, g)R(f, g_0^*)$ if $r=0$, $s=m$, and a similar remark holds in the case $r=n$, $s=0$. Hence we have proved that in all cases

$$(2)\qquad R_{10} = c\cdot R(f_r, g_s)R(f^*_{n-r}, g^*_{m-s}),$$

where c is an integer.

The resultant $R(\bar{f}, \bar{g})$ can be obtained from $R(f, g)$ by replacing $a_0, a_1, \cdots, a_{r-1}$ and $b_0, b_1, \cdots, b_{s-1}$ by $a_0t^r, a_1t^{r-1}, \cdots, a_{r-1}t$ and $b_0t^s, b_1t^{s-1}, \cdots, b_{s-1}t$ respectively. Every term of $R(f, g)$ acquires then a factor t^w, where w is the weight of this term as specified in the statement of Theorem 1. By (1), $R_{10}(=R_1(a_i, b_j, 0))$ is the sum of all terms of $R(f, g)$ of lowest weight k, and since, always according to our definition of the weight, $R(f_r, g_s)$ is isobaric of weight rs, while $R(f^*_{-r}, g^*_{m-s})$ is of weight zero, it follows that $k=rs$. This and the identity (2) complete the proof of our theorem.

To determine the numerical constant c, we take a special case, say $f=a_0y^n+a_n$, $g=b_sy^{m-s}$. Then $f_r=a_0y^r$, $g_s=b_s$, $f^*_{n-r}=a_n$, $g^*_{m-s}=b_sy^{m-s}$, and

$$R(f, g) = (-1)^{(m-s)n}a_0^s a_n^{m-s}b_s^n,$$
$$R(f_r, g_s) = a_0^s b_s^r,\quad R(f^*_{n-r}, g^*_{m-s}) = (-1)^{(n-r)(m-s)}a_n^{m-s}b_s^{n-r}.$$

Hence, in this case we have

$$R(f, g) = (-1)^{(m-s)r}R(f_r, g_s)R(f^*_{n-r}, g^*_{m-s}),$$

and consequently $c=(-1)^{(m-s)r}$.

Remark. The resultant of the polynomials f and g coincides, to within the sign, with the resultant of the polynomials $a_ny^n+\cdots+a_0$, $b_my^m+\cdots+b_0$. Applying our theorem to the last two polynomials, we see that it is permissible to interchange, in the statement of Theorem 1, a_i with a_{n-i} and b_j with b_{n-j}. This is equivalent to attaching the weights $r, r-1, \cdots, 1, s, s-1, \cdots, 1$ to $a_n, a_{n-1}, \cdots, a_{n-r+1}, b_m, b_{m-1}, \cdots, b_{m-s+1}$ respectively, and the weight 0 to the remaining coefficients.

3. **The intersection multiplicity of two curves at a common point.** The application of Theorem 1 toward the determination of the intersection multiplicity of two curves at a common point is immediate. If the coefficients a_i and b_j of the polynomials f and g are polynomials in x, and if the origin O is a common point of the two curves $f=f(x, y)=0$, and $g=g(x, y)=0$,

r-fold for f and s-fold for g, then $a_n, a_{n-1}, \cdots, a_{n-r+1}$ are divisible by x^r, $x^{r-1}, \cdots, x$, respectively and $b_m, b_{m-1}, \cdots, b_{m-s+1}$ are divisible by $x^s, x^{s-1}, \cdots, x$, respectively. Hence every term of the resultant $R(f, g) = R(x)$ is divisible by x^w, where w is the weight of the term as specified in the remark at the end of the preceding section. Since $w \geqq rs$, x^{rs} divides $R(x)$. Let

$$R(x) = \alpha x^{rs} + \text{terms of higher degree},$$

where α is a constant.

Let $f = \sum c_{ij} x^i y^j$, $g = \sum d_{ij} x^i y^j$. Then

$$c_{ij} = \left[\frac{a_{n-j}(x)}{x^{r-j}}\right]_{x=0},$$

for all i and j such that $i+j=r$; similarly

$$d_{ij} = \left[\frac{b_{m-j}(x)}{x^{s-j}}\right]_{x=0},$$

if $i+j=s$. Moreover, $c_{0,j} = a_{n-j}(0)$, $j=r, r+1, \cdots, n$, and $d_{0,j} = b_{m-j}(0)$, $j=s, s+1, \cdots, m$. Applying Theorem 1, we find

$$\alpha = \pm R(f_r, g_s) R(f^*_{n-r}, g^*_{m-s}),$$

where

$$f_r = c_{r0}x^r + c_{r-1,1}x^{r-1}y + \cdots + c_{0r}y^r,$$
$$g_s = d_{s0}x^s + d_{s-1,1}x^{s-1}y + \cdots + d_{0s}y^s,$$

and

$$f^*_{n-r} = c_{0r} + c_{0,r+1}y + \cdots + c_{0,n}y^{n-r},$$
$$g^*_{m-s} = d_{0s} + d_{0,s+1}y + \cdots + d_{0,m}y^{m-s}.$$

Here $R(f_r, g_s) = 0$ if and only if the curves f and g have at the origin a common principal tangent. If $R(f_r, g_s) \neq 0$, then $R(f^*_{n-r}, g^*_{m-s}) = 0$ if and only if the two curves have a common point on the y-axis outside the origin. Hence if the y-axis is generic and if there are no common principal tangents at O, then $\alpha \neq 0$ and the intersection multiplicity at O equals rs.

II. The general case of $n+1$ polynomials in n indeterminates

4. Preliminary remarks on forms of inertia. Let K be an underlying domain of integrity, and let $f_1, f_2, \cdots, f_m$ be polynomials in $x_1, x_2, \cdots, x_n$, with coefficients in a polynomial ring $K[t] = K[t_1, t_2, \cdots, t_s]$, where $t_1, \cdots, t_s$ are indeterminates. A polynomial T in $K[t]$ is an *inertia form* of the polynomials $f_1, \cdots, f_m$, if it has the property:

$$(3) \qquad x_i^{\tau} T \equiv 0(f_1, \cdots, f_m),$$

for $i=1, 2, \cdots, n$ and for some integer τ, i.e., if $x_i^\tau T$ belongs to the polynomial ideal generated by $f_1, \cdots, f_m$ in $K[t_1, \cdots, t_s; x_1, \cdots, x_n]$. It follows from the definition that the inertia forms of $f_1, \cdots, f_m$ form an ideal $\mathfrak{T}$ in $K[t]$.

THEOREM 2. *If for $\alpha=1, 2, \cdots, n$ each polynomial f_j is of the form: $f_j = t_{\alpha_j} x_\alpha^{\sigma_{j\alpha}} + f_{j\alpha}^*$, $\sigma_{j\alpha} \geqq 0$, where $t_{\alpha_1}, \cdots, t_{\alpha_m}$ are distinct indeterminates in the set $t_1, \cdots, t_s$ and where $f_{j\alpha}^*$ is a polynomial independent of $t_{\alpha_1}, \cdots, t_{\alpha_m}$, then $\mathfrak{T}$ is a prime ideal, and* (3) *holds for $i=1, 2, \cdots, n$ if it holds for one value of i.*

In order to prove† the theorem let, for instance, $f_j = t_j x_1^{\sigma_j} + f_j^*$. If (3) holds for a given i and for a given polynomial $T(t_1, t_2, \cdots, t_m, \cdots, t_s)$, then it follows by the substitution $t_j = -f_j^*/x_1^{\sigma_j}$:

$$T\left(-\frac{f_1^*}{x_1^{\sigma_1}}, -\frac{f_2^*}{x_1^{\sigma_2}}, \cdots, -\frac{f_m^*}{x_1^{\sigma_m}}, \cdots, t_s\right) = 0. \tag{4}$$

Conversely, if a polynomial $T(t_1, \cdots, t_s)$ vanishes identically after the substitution $t_j = -f_j^*/x_1^{\sigma_j}$, then T satisfies (3) for $i=1$. Under the assumption made in the above theorem, it follows immediately that (4) is a necessary and sufficient condition in order that T be a form of inertia. Hence $\mathfrak{T}$ is a prime ideal and T is a form of inertia if (3) holds for $i=1$.

COROLLARY. *If $\sigma_1 = \sigma_2 = \cdots = \sigma_m = 0$, then any form of inertia T satisfies* (3) *with $\tau = 0$.*

If the polynomials $f_1, f_2, \cdots, f_m$ are homogeneous in $x_1, \cdots, x_n$, then it is well known that the vanishing of all the inertia forms for special values t_j^0 of the parameters t_j is a necessary and sufficient condition that the equations $f_1(x_i; t_j) = 0, f_2(x_i; t_j^0) = 0, \cdots, f_m(x_i; t_j^0) = 0$ have a non-trivial solution (not all $x_i = 0$) (see W., p. 16).

For non-homogeneous polynomials the following theorem holds:

THEOREM 3.1. *Let f_i contain terms of lowest degree s_i in $x_1, \cdots, x_n$:*

$$f_i = f_{i,s_i}(x_1, \cdots, x_n) + f_{i,s_i+1}(x_1, \cdots, x_n) + \cdots,$$

where $f_{i,k}$ is homogeneous of degree k in $x_1, \cdots, x_n$, and let us consider the homogeneous polynomials:

$$\bar{f}_i = x_0^{l_i - s_i} f_{i,s_i}(x_1, \cdots, x_n) + x_0^{l_i - s_i - 1} f_{i,s_i+1}(x_1, \cdots, x_n) + \cdots, \tag{5}$$

where x_0 is an indeterminate and l_i is the degree of f_i. The vanishing of all the inertia forms of $f_1, f_2, \cdots, f_m$ for special values of the parameters t_j is a sufficient condition in order that (a) *either the equations $\bar{f}_1 = 0, \cdots, \bar{f}_m = 0$ have a non-*

† Compare W., p. 15.

trivial solution different from $x_0=1$, $x_1=\cdots=x_n=0$; *or that* (b) *the equations* $f_{1,s_1}(x_1,\cdots,x_n)=0,\cdots,f_{m,s_m}(x_1,\cdots,x_n)=0$ *have a non-trivial solution (in a suitable extension field of* K).

The converse holds only under certain restrictions:

THEOREM 3.2. *If* (a) *holds and if the coefficients of* $x_1^{l_i},\cdots,x_n^{l_i}$ *in* f_i $(i=1,2,\cdots,m)$ *are indeterminates which do not occur in other terms of* f_i, *then the inertia forms of* $f_1,\cdots,f_m$ *all vanish.*

THEOREM 3.3. *If* (b) *holds, and if the coefficients of* $x_1^{s_i},\cdots,x_n^{s_i}$ *in* f_i $(i=1,2,\cdots,m)$ *are indeterminates which do not occur in other terms of* f_i, *then the inertia forms of* $f_1,\cdots,f_m$ *all vanish.*

$\bar{f}_1,\cdots,\bar{f}_m$, considered as polynomials in x_0, possess a resultant system

$$\phi_1(x_1,\cdots,x_n),\cdots,\phi_h(x_1,\cdots,x_n),$$

where the ϕ_i's are homogeneous polynomials. Since $\phi_i\equiv 0(\bar{f}_1,\cdots,\bar{f}_m)$, we have for every inertia form T of the polynomials ϕ_i: $x_j^{\sigma}T\equiv 0(\bar{f}_1,\cdots,\bar{f}_m)$, $j=1,2,\cdots,n$. Putting $x_0=1$, we see that T is also an inertia form of the polynomials $f_1,\cdots,f_m$.

Let all the inertia forms of $f_1,\cdots,f_m$ vanish for special values of the parameters t_j. Then for these special values of the t_j's also the inertia forms of $\phi_1,\cdots,\phi_h$ all vanish, the *homogeneous equations* $\phi_1=0,\cdots,\phi_h=0$ have a non-trivial solution, and consequently, by known properties of the resultant system $\phi_1,\cdots,\phi_h$, the alternatives (a) and (b) of Theorem 3.1 follow.

If T is an inertia form of $f_1,\cdots,f_m$, then passing to the homogeneous polynomials $\bar{f}_1,\cdots,\bar{f}_m$, it is found that $(x_0x_i)^{\sigma}T\equiv 0(\bar{f}_1,\cdots,\bar{f}_m)$, for $i=1,2,\cdots,n$ and for some σ. Under the hypothesis of Theorem 3.2 concerning the coefficients of $x_1^{l_i},\cdots,x_n^{l_i}$, we can repeat the reasoning of the proof of Theorem 2, and it follows that $x_i^{\rho}T\equiv 0(\bar{f}_1,\cdots,\bar{f}_m)$, for $i=1,2,\cdots,n$ and for some ρ. Hence if (a) holds, then $T=0$.

For the proof of Theorem 3.3, let $x_1^0,\cdots,x_n^0$ be a non-trivial solution of the equations $f_{1,s_1}=0,\cdots,f_{m,s_m}=0$, and let, for instance, $x_1^0\neq 0$. We make the following change of indeterminates:

$$x_1=y_1,\ x_2=y_2y_1,\cdots,x_n=y_ny_1.$$

Then

$$\begin{aligned} f_i &= y_1^{s_i}f_{1,s_i}(1,y_2,\cdots,y_n)+y_1^{s_i+1}f_{1,s_i+1}(1,y_2,\cdots,y_n)+\cdots \\ &= y_1^{s_i}\psi_i(y_1,\cdots,y_n), \end{aligned}$$

and if T is an inertia form of $f_1(x),\cdots,f_m(x)$, then $y_1^{\sigma}T\equiv 0(\psi_1,\cdots,\psi_m)$. Under the hypothesis of Theorem 3.3, the constant terms in $\psi_1,\psi_2,\cdots,\psi_m$ are in-

determinates, and hence, by the corollary to Theorem 2, $T\equiv 0(\psi_1, \cdots, \psi_m)$. Since for $t_i = t_i^0$, the equations $\psi_1 = 0, \cdots, \psi_m = 0$ have the solution $y_1^0 = 0$, $y_2^0 = x_2^0/x_1^0, \cdots, y_n = x_n^0/x_1^0$, it follows that $T(t_1^0, \cdots, t_s^0) = 0$.

5. **The inertia forms of some special set of $n+1$ polynomials in n indeterminates.** The theorems of the preceding section are applicable in the special case when $m = n+1$ and when each f_i is a polynomial with literal coefficients in which all the terms of degree $< s_i \leqq l_i$ are missing, l_i being the degree of f_i:

$$(6) \qquad f_i = \sum_{(j)} a^{(i)}_{j_1 j_2 \cdots j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}, \qquad s_i \leqq j_1 + \cdots + j_n \leqq l_i.$$

If $s_1 = s_2 = \cdots = s_{n+1} = 0$, then the ideal of the inertia forms is a principal ideal (R), where R is the resultant of $f_1, f_2, \cdots, f_{n+1}$. R is an irreducible polynomial homogeneous of degree $l_2 \cdots l_{n+1}$ in the coefficients of f_1, homogeneous of degree $l_1 l_3 \cdots l_{n+1}$ in the coefficients of f_2, etc. Finally, by the corollary to Theorem 2, $R\equiv 0(f_1, f_2, \cdots, f_{n+1})$, and the vanishing of R for special values of the coefficients $a^{(i)}_{(j)}$ is a necessary and sufficient condition in order that the polynomials $f_1, f_2, \cdots, f_{n+1}$, rendered homogeneous, have a common non-trivial zero (see W., p. 20).

We prove the following theorems in the case when $s_1, s_2, \cdots, s_{n+1}$ are not necessarily all zero:

THEOREM 4. *Let e_i $(= a^{(i)}_{l_i 0 \cdots 0})$ be the coefficient of $x_1^{l_i}$ in f_i. If $s_{n+1} < l_{n+1}$, then any inertia form of $f_1, f_2, \cdots, f_{n+1}$ which does not vanish identically, must be of degree >0 in each of the coefficients $e_1, e_2, \cdots, e_n$.*

COROLLARY. *If one at least of the polynomials $f_1, \cdots, f_{n+1}$ is non-homogeneous, the ideal $\mathfrak{T}$ of their inertia forms is a principal ideal.*†

The proof is similar to the one given in W., pp. 16–17, in the case $s_1 = \cdots = s_{n+1} = 0$, only with a slightly different specialization of the coefficients $a^{(i)}_{(j)}$. Assume that there exists an inertia form T, not identically zero, which is independent of e_1. Putting $f_i = e_i x_1^{l_i} + f_i^*$, and applying (4) (where σ_i should be replaced by l_i), we see that T cannot be independent of all the coefficients $e_2, \cdots, e_{n+1}$ (since T is not identically zero) and we conclude that the quotients

$$f_2^*/x_1^{l_2}, \cdots, f_{n+1}^*/x_1^{l_{n+1}}$$

are algebraically dependent in $K[a^{(i)}_{(j)}]$, K being the ring of natural integers. By a lemma proved in W., p. 17, these quotients remain algebraically depend-

† If all the polynomials f_i are homogeneous, then $\mathfrak{T}$ contains the resultant of any n of these polynomials and is therefore not a principal ideal.

ent after an arbitrary specialization $a_{(j)}^{(i)}=\alpha_{(j)}^{(i)}$ $(\alpha_{(j)}^{(i)}<K)$. Let us take for $f_1, f_2, \cdots, f_{n+1}$ the special set of polynomials $x_1^{l_1}, x_1^{l_2-1}x_2, \cdots, x_1^{l_n-1}x_n, x_1^{l_{n+1}-1}$, observing that the specialization $f_{n+1}=x_1^{l_{n+1}-1}$ is permissible, since, by hypothesis, f_{n+1} is not homogeneous. The above quotients become

$$x_2/x_1, \cdots, x_n/x_1, 1/x_1,$$

and since these are evidently algebraically independent, our assumption that T is independent of e_1 leads to a contradiction.

The corollary now follows in exactly the same manner as in W., p. 19.

Let (D) be the principal ideal of the inertia forms of the polynomials $f_1, f_2, \cdots, f_{n+1}$. D, if it is not identically zero, is an irreducible polynomial in the coefficients $a_{(j)}^{(i)}$. We next prove that indeed *D is not identically zero*, i.e., that there exist inertia forms of $f_1, \cdots, f_{n+1}$ which are not identically zero.

If $\phi_1, \cdots, \phi_{n+1}$ denote general polynomials in $x_1, \cdots, x_n$ with literal coefficients, of degree $l_1, l_2, \cdots, l_{n+1}$ respectively, we can write $\phi_i=\psi_i+f_i$, where $f_1, \cdots, f_{n+1}$ are our given polynomials and where ψ_i is of degree s_i-1. Let $\phi_i=\sum a_{j_1\cdots j_n}^{(i)}x_1^{j_1}\cdots x_n^{j_n}$, $0\leqq j_1+\cdots+j_n\leqq l_i$. Let t be a parameter, and let ϕ_i^t be the polynomial obtained from ϕ_i by replacing each coefficient $a_{j_1\cdots j_n}^{(i)}$ by $t^{s_i-j_1-\cdots-j_n}a_{j_1\cdots j_n}^{(i)}$, if $s_i>j_1+\cdots+j_n$, i.e., if $a_{j_1\cdots j_n}^{(i)}$ is the coefficient of a term of the polynomial ψ_i, while the coefficients of f_i remain unaltered. Let $R_t=R(\phi_1^t, \cdots, \phi_{n+1}^t)$ be the resultant of the ϕ_i^t's considered as polynomials in $x_1, \cdots, x_n$, and let t^α, $\alpha\geqq 0$, be the highest power of t which divides R_t:

$$R_t = t^\alpha R^{(1)}(t, a_{(j)}^{(i)}) = t^\alpha R_t^{(1)}. \tag{7}$$

Since each polynomial ϕ_i^t contains the terms $x_1^{l_i}, \cdots, x_n^{l_i}$, whose coefficients are indeterminates, it follows by Theorem 2, that the ideal of the inertia forms of $\phi_1^t, \cdots, \phi_{n+1}^t$ is prime. Now, no power of t is an inertia form of $\phi_1^t, \cdots, \phi_{n+1}^t$, because otherwise, for $t=1$, it would follow that 1 is an inertia form of $\phi_1, \cdots, \phi_{n+1}$, and this is impossible. Hence, since $t^\alpha R_t^{(1)}$ is an inertia form of $\phi_1^t, \cdots, \phi_{n+1}^t$, it follows that also $R_t^{(1)}$ is an inertia form. For $t=0$, we have $\phi_i^0=f_i$, and $R_0^{(1)}$ is therefore an inertia form of $f_1, \cdots, f_{n+1}$ which does not vanish identically.

6. The resultant $R(\phi_1, \cdots, \phi_{n+1})$ as an isobaric function of the coefficients $a_{(j)}^{(i)}$. Let $\phi_1, \cdots, \phi_{n+1}$ denote, as in the preceding section, general polynomials in the n variables $x_1, \cdots, x_n$, of degree $l_1, \cdots, l_{n+1}$ respectively, and let $R(\phi_1, \cdots, \phi_{n+1})=R(a_{(j)}^{(i)})$ be their resultant. It is clear that $R(\cdots, t^{j_1+\cdots+j_n}a_{j_1\cdots j_n}^{(i)}, \cdots)$ is the resultant of $\phi_1(x_1t, \cdots, x_{n+1}t), \cdots, \phi_{n+1}(x_1t, \cdots, x_{n+1}t)$ and is therefore divisible by $R(a_j^{(i)})$, since the ideal of the inertia forms of these $n+1$ polynomials is, by the preceding theorems,

a principal ideal and since the irreducible polynomial $R(a^{(i)}_{(j)})$ obviously belongs to this ideal. It follows that $R(\cdots, t^{j_1+\cdots+j_n}a^{(i)}_{j_1\cdots j_n}, \cdots)$ differs from $R(a_j^{(i)})$ only by a factor which is a power of t, say by t^σ. Hence *$R(a_j^{(i)})$ is an isobaric function of the coefficients of $a_j^{(i)}$, of weight σ, provided that we attach to $a^{(i)}_{j_1\cdots j_n}$ the weight $j_1+\cdots+j_n$.*

To find σ, we specialize the polynomials ϕ_i as follows

$$\phi_1 = a_1x_1^{l_1}, \cdots, \phi_n = a_nx_n^{l_n}, \phi_{n+1} = a_{n+1}.$$

The resultant R does not vanish identically, since the equations $\phi_1=0, \cdots, \phi_{n+1}=0$ have no common solution if $a_1, \cdots, a_{n+1}$ are indeterminates. Taking into account the degree of R in the coefficients of each ϕ_i, we deduce that $R=c\cdot a_1^{l_2\cdots l_{n+1}}\cdots a_{n+1}^{l_1l_2\cdots l_n}$, where c is a numerical factor. Since $a_1, \cdots, a_n$ are of weight $l_1, l_2, \cdots, l_n$ respectively and a_{n+1} is of weight zero, it follows that $\sigma=nl_1l_2\cdots l_{n+1}$.

As an immediate corollary of this last result and of the fact that $R(\phi_1, \cdots, \phi_{n+1})$ is homogeneous of degree $l_1\cdots l_{i-1}l_{i+1}\cdots l_{n+1}$ in the coefficients of ϕ_i, it follows that *if we attach to $a^{(i)}_{j_1\cdots j_n}$ the weight $l_i-j_1-\cdots-j_{n+1}$, then $R(\phi_1, \cdots, \phi_{n+1})$ is isobaric of weight $l_1l_2\cdots l_{n+1}$.*

7. **Properties of R based on a more general definition of the weights of the coefficients $a^{(i)}_{(j)}$.** We separate in ϕ_i the terms of degree $\leqq s_i$ from those of degree $>s_i$, and we put $\phi_i=\bar{\psi}_i+\bar{f}_i$, where $\bar{\psi}_i$ is of degree s_i and $\bar{f}_i$ contains all the terms of degree $>s_i$. While in §5 we have replaced $a^{(i)}_{j_1\cdots j_n}$ by $t^{s_i-j_1-\cdots-j_n}a^{(i)}_{j_1\cdots j_n}$, if $s_i>j_1-\cdots-j_n$, we now instead replace $a^{(i)}_{j_1\cdots j_n}$ by $t^{j_1+\cdots+j_n-s_i}a^{(i)}_{j_1\cdots j_n}$, if $j_1+\cdots+j_n>s_i$, i.e., if $a^{(i)}_{j_1\cdots j_n}$ is the coefficient of a term in $\bar{f}_i$, and leave the coefficients of $\bar{\psi}_1, \cdots, \bar{\psi}_{n+1}$ unaltered.

Let $\bar{\phi}_1^t, \cdots, \bar{\phi}_{n+1}^t$ be the polynomials obtained in this manner, and let

$$R(\bar{\phi}_1^t, \cdots, \bar{\phi}_{n+1}^t) = t^\beta R_t^{(2)} = t^\beta R^{(2)}(t; a^{(i)}_{(j)}) \tag{8}$$

be the resultant of the polynomials $\bar{\phi}_i^t$. Here t^β is the highest power of t which divides $R(\bar{\phi}_1^t, \cdots, \bar{\phi}_{n+1}^t)$, so that $R_0^{(2)}=R^{(2)}(0; a_j^{(i)})$ does not vanish identically. As in the case of the polynomials ϕ_i^t of §5, we conclude also here that $R_0^{(2)}$ is a form of inertia of the polynomials $\bar{\psi}_1, \cdots, \bar{\psi}_{n+1}$, and since these are general polynomials of degree $s_1, \cdots, s_{n+1}$ respectively, we deduce that *$R_0^{(2)}$ is divisible by $R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$.*

Now the polynomials ϕ_i^t and $\bar{\phi}_i^t$ are related in the following way: $\bar{\phi}_i^t=\phi_i^t(tx_1, \cdots, tx_{n+1})/t^{s_i}$. From this it follows, in view of the isobaric property of R given in the preceding section, that their resultants differ only by a factor which is a power of t. Hence, by (7) and (8), we have $R^{(1)}(t, a^{(i)}_{(j)})=R^{(2)}(t, a^{(i)}_{(j)})$, and in particular for $t=0$, we have $R_0^{(1)}=R_0^{(2)}$. Let $R_0=R_0^{(1)}=R_0^{(2)}$. R_0 is divisible by $R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$ *and* by D, where D is the base of

the principal ideal of the inertia forms of $f_1, \cdots, f_{n+1}$.† Hence R_0 *is divisible by the product* $D \times R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$, since both factors are irreducible and distinct polynomials. ($R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$ is of degree >0 in each constant term $a^{(1)}_{00\cdots 0}$, while, except in the trivial case $s_1 = \cdots = s_{n+1} = 0$, where $f_1, \cdots, f_{n+1}$ coincide with $\phi_1, \cdots, \phi_{n+1}$, at least one of the polynomials f, say f_i, and hence also D, is independent of $a^{(1)}_{00\cdots 0}$.)

The precise relationship between R_0 and $D \cdot R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$ is given by the following theorems:

THEOREM 5.1. *If two at least of the polynomials f_i are non-homogeneous, then*

$$R_0 = c \cdot D \cdot R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1}), \tag{9.1}$$

where c is a numerical factor (an integer).

THEOREM 5.2. *If $f_2, \cdots, f_{n+1}$ are homogeneous, then*

$$R_0 = c \cdot D^{l_1 - s_1} \cdot R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1}), \tag{9.2}$$

where c is a numerical factor (an integer). In this case D is simply the resultant of $f_2, \cdots, f_{n+1}$.

Before proving these theorems, let us first derive an immediate consequence. From the meaning of $R_0 = R_0^{(1)}$ [cf. (7)] it follows that if to each coefficient $a^{(i)}_{j_1 \cdots j_n}$ in ϕ_i we attach the weight $s_i - j_1 - \cdots - j_n$, if $j_1 + \cdots + j_n \leqq s_i$, and the weight zero if $j_1 + \cdots + j_n > s_i$, then R_0 is the sum of terms of lowest weight α in the resultant $R(\phi_1, \cdots, \phi_{n+1})$. According to this definition of the weight, each term in D is of weight zero, while $R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$, by §6, is of weight $s_1 \cdots s_{n+1}$. Hence we may state the following theorem:

THEOREM 6. *Let $\phi_1, \cdots, \phi_{n+1}$ be general polynomials in $x_1 \cdots x_n$, of degree $l_1, \cdots, l_{n+1}$ respectively, and let $s_1, \cdots, s_{n+1}$ be integers such that $0 \leqq s_i \leqq l_i$. If we attach to each coefficient $a^{(i)}_{j_1 \cdots j_n}$ in ϕ_i the weight $s_i - j_1 - \cdots - j_n$ or the weight zero, according as $j_1 + \cdots + j_n \leqq s_i$ or $j_1 + \cdots + j_n > s_i$, then each term of the resultant $R(\phi_1, \cdots, \phi_{n+1})$ is of weight $\geqq s_1 s_2 \cdots s_{n+1}$. The sum of terms of lowest weight $s_1 s_2 \cdots s_{n+1}$ is given by the product $cD^\sigma R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$, where c is a numerical factor. The symbols have the following meaning: $\bar{\psi}_i$ is the sum of terms of ϕ_i which are of degree $\leqq s_i$ and f_i is the sum of terms of ϕ_i of degree $\geqq s_i$; $R(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$ is the resultant of $\bar{\psi}_1, \cdots, \bar{\psi}_{n+1}$; if not all $s_i = l_i$, then D is the base of the principal ideal of the inertia forms of $f_1, \cdots, f_{n+1}$; if all $s_i = l_i$, then $D = 1$; finally, $\sigma = 1$, except when all the integers s_i but one, say s_1, coincide with the corresponding integers l_i, in which case $\sigma = l_1 - s_1$.*

---

† In the trivial case when $f_1, \cdots, f_{n+1}$ are all homogeneous polynomials, D is not defined, but then R_0 evidently coincides with $R(\phi_1, \cdots, \phi_{n+1})$.

Remark. Again from the meaning of $R_0(=R_0^{(2)})$ it follows that if we attach to $a^{(i)}_{j_1\cdots j_n}$ the weight $j_1+\cdots+j_n-s_i$ or zero, according as $j_1+\cdots+j_n\geqq s_i$ or $j_1+\cdots+j_n<s_i$, then R_0 is also the sum of terms of lowest weight, β, in the resultant $R(\phi_1,\cdots,\phi_{n+1})$ [cf. (8)]. According to this definition of the weight, each term of $R(\bar\psi_1,\cdots,\bar\psi_{n+1})$ is of weight zero, and D^σ has to be isobaric of weight β.

To find β, we observe that Theorem 5.1 implies that D is homogeneous of degree $l_2\cdots l_{n+1}-s_2\cdots s_{n+1}$ in the coefficients of f_1, homogeneous of degree $l_1l_3\cdots l_{n+1}-s_1s_3\cdots s_{n+1}$ in the coefficients of f_2, etc. On the other hand, if $j_1+\cdots+j_n$ is taken as the weight of $a^{(i)}_{j_1\cdots j_n}$, then R_0 and $R(\bar\psi_1,\cdots,\bar\psi_{n+1})$ are isobaric forms of weight $nl_1\cdots l_{n+1}$ and $ns_1\cdots s_{n+1}$ respectively, whence D is of weight $n(l_1\cdots l_{n+1}-s_1\cdots s_{n+1})$. It follows that if we replace in the polynomial D each coefficient $a^{(i)}_{j_1\cdots j_n}$ by $a^{(i)}_{j_1\cdots j_n}t^{s_i-j_1-\cdots-j_n}$, D acquires the factor t^β, where

$$\begin{aligned}\beta = &-n(l_1\cdots l_{n+1}-s_1\cdots s_{n+1})+s_1(l_2\cdots l_{n+1}-s_2\cdots s_{n+1})\\ &+s_2(l_1l_3\cdots l_{n+1}-s_1s_3\cdots s_{n+1})+\cdots+s_{n+1}(l_1\cdots l_n-s_1\cdots s_n),\end{aligned}$$

or

$$\beta = s_1s_2\cdots s_{n+1}+\left(\frac{l_1-s_1}{l_1}+\cdots+\frac{l_{n+1}-s_{n+1}}{l_{n+1}}-1\right)l_1l_2\cdots l_{n+1}.$$

If $s_2=l_2,\cdots,s_{n+1}=l_{n+1}$, then it is seen that $\beta=0$, and this agrees with Theorem 5.2, because in this case the coefficients of $f_2,\cdots,f_{n+1}$ are of weight zero.

Proof of Theorems 5.1 and 5.2. We begin with Theorem 5.2, whose proof is simpler. We have in this case $\bar\psi_i=\phi_i$, $i=2,\cdots,n+1$, and hence $R(\bar\psi_1,\cdots,\bar\psi_{n+1})$ is of degree $l_2\cdots l_{n+1}$ in the coefficients of $\bar\psi_1$. Hence, if we put

$$R_0 = D^\sigma R(\bar\psi_1,\phi_2,\cdots,\phi_{n+1})\cdot P,$$

then P is independent of the coefficients of ϕ_1.

Now in the present case $\beta=0$, and R_0 is what becomes of the resultant $R(\bar\phi_1^t,\cdots,\bar\phi_{n+1}^t)$ if we put $t=0$, where now $\phi_i^t=\phi_i$, $i=2,\cdots,n+1$, and $\bar\phi_1^t=\bar\psi_1+\bar f_1(tx_1,\cdots,tx_n)/t^{s_1}$. It follows that $R_0=0$ implies that either the equations $\bar\psi_1=0$, $\phi_2=0,\cdots,\phi_{n+1}=0$, rendered homogeneous, have a non-trivial solution, or that the homogeneous equations $f_2=0,\cdots,f_{n+1}=0$ have a non-trivial solution. Hence $R(\bar\psi_1,\phi_2,\cdots,\phi_{n+1})$ and $R(f_2,\cdots,f_{n+1})$ are the only irreducible factors which can occur in R_0. Since the *irreducible* poly-

nomial $R(f_2, \cdots, f_{n+1})$ obviously coincides with D, Theorem 5.2 follows by comparing the degrees of the first and second member of (9.2).

For the proof of Theorem 5.1, it is sufficient to show that *D is of degree $l_2l_3 \cdots l_{n+1} - s_2s_3 \cdots s_{n+1}$ in the coefficients of f_1, of degree $l_1l_3 \cdots l_{n+1} - s_1s_3 \cdots s_{n+1}$ in the coefficients of f_2*, etc. Since $DR(\bar{\psi}_1, \cdots, \bar{\psi}_{n+1})$ divides R_0, D cannot be of higher degree in the coefficients of $f_1, f_2, \cdots, f_{n+1}$, and therefore it remains to show that D is of degree *not less* than $l_2l_3 \cdots l_{n+1} - s_2s_3 \cdots s_{n+1}$ in the coefficients of f_1, etc. We prove this in the following section.

8. **The degree of D.** We wish to show in this section that *if at least two of the polynomials $f_1, \cdots, f_{n+1}$ are non-homogeneous, then D is of degree $\geqq l_2 \cdots l_{n+1} - s_2 \cdots s_{n+1}$ in the coefficients of f_1, of degree $\geqq l_1l_3 \cdots l_{n+1} - s_1s_3 \cdots s_{n+1}$ in the coefficients of f_2*, etc. Obviously, the condition that at least two of the polynomials f_i be non-homogeneous, is necessary. In fact, if only one of the polynomials f_i, say f_{n+1}, is non-homogeneous, then D coincides with the resultant $R(f_1, \cdots, f_n)$ of the forms $f_1, \cdots, f_n$, and its degree in the coefficients of f_1 is not $l_2 \cdots l_{n+1} - s_2 \cdots s_{n+1}$ $[= l_2 \cdots l_n(l_{n+1} - s_{n+1})]$, but $l_2 \cdots l_n$. If all the polynomials f_i are homogeneous, then the ideal of their inertia forms is not a principal ideal and D is not defined.

If for special values of the coefficients $a^{(i)}_{(j)}$, one of the polynomials f_i, say f_{n+1}, factors into a product gh of two polynomials, then D becomes an inertia form of both sets of polynomials $f_1, \cdots, f_n, g$ and $f_1, \cdots, f_n, h$. Hence, assuming that the ideals of inertia forms of these two sets of polynomials are principal ideals, say (D_1) and (D_2) respectively, then for those special values of the coefficients $a^{(i)}_{(j)}$, D is divisible by both D_1 and D_2. This remark shall be used in the sequel.

Let f_n and f_{n+1} be the non-homogeneous polynomials. We first consider the case in which $f_1, \cdots, f_{n-1}$ are polynomials of degree 1, and in this case we examine separately three possibilities.

(a) *At least two of the polynomials $f_1, \cdots, f_{n-1}$ are non-homogeneous* (and hence two at least of the integers $s_1, \cdots, s_{n-1}$ vanish). We specialize the coefficients of f_n and f_{n+1} in such a manner that f_n becomes the product of l_n general polynomials $f_{n,i}$ of the first degree, of which s_n are linear forms, and that f_{n+1} becomes similarly the product of l_{n+1} linear factors, $f_{n+1,i}$. The l_nl_{n+1} $(n+1)$-row coefficient determinants relative to the sets of polynomials $f_1, \cdots, f_{n-1}, f_{n,i}, f_{n+1,j}$ are all distinct and irreducible inertia forms, since at least two of the polynomials of each set are non-homogeneous. Hence D is divisible by the product of these determinants and is therefore of degree $\geqq l_nl_{n+1}$ in the coefficients of f_i, $i = 1, 2, \cdots, n-1$, and of degree $\geqq l_{n+1}$ (l_n) in the coefficients of f_n (f_{n+1}).

We observe that this proves that in the present case D coincides with the

resultant $R(f_1, \cdots, f_{n+1})$, or, what is the same, that this resultant is irreducible.

(b) *All but one of the polynomials* $f_1, \cdots, f_{n-1}$ *are homogeneous*. Let, for instance, f_1 be non-homogeneous. With the same specialization of f_n and f_{n+1} as in the preceding case, let $f_{n,1}, \cdots, f_{n,s_n}; f_{n+1,1}, \cdots, f_{n+1,s_{n+1}}$ be the homogeneous linear factors of f_n and of f_{n+1} respectively. The $(n+1)$-row coefficient determinants of $f_1, \cdots, f_{n-1}, f_{n,i}, f_{n+1,j}$ remain irreducible, except when simultaneously $1 \leqq i \leqq s_n$ *and* $1 \leqq j \leqq s_{n+1}$, in which case the determinant factors into the constant term of f_1 and into the n-row determinant of the coefficients of $x_1, \cdots, x_n$ in $f_2, \cdots, f_{n-1}, f_{n,i}, f_{n+1,j}$. Hence D is divisible by the product of $l_n l_{n+1} - s_n s_{n+1}$ $(n+1)$-row determinants and $s_n s_{n+1}$ n-row determinants, these last ones being independent of the coefficients of f_1. Hence D is of degree $\geqq l_n l_{n+1} - s_n s_{n+1}$ in the coefficients of f_1, of degree $\geqq l_n l_{n+1}$ in the coefficients of f_i, $i = 2, \cdots, n-1$, and of degree $\geqq l_{n+1}$ (l_n) in the coefficients of f_n (f_{n+1}).

(c) *All the polynomials* $f_1, \cdots, f_{n-1}$ *are homogeneous*. Let

$$f_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, 2, \cdots, n-1, \tag{10}$$

$$f_i = f_{i,s_i} + f_{i,s_i+1} + \cdots + f_{i,l_i}, \qquad i = n, n+1, \tag{10'}$$

where f_{i,s_i+k} is homogeneous of degree $s_i + k$. Solving (10) for $x_2, \cdots, x_n$ we get

$$A_1 x_i \equiv A_i x_1(f_1, f_2, \cdots, f_{n-1}), \tag{11}$$

where $A_1, \cdots, A_n$ are $(n-1)$-row minors of the matrix (a_{ij}) and hence homogeneous of degree 1 in the coefficients of each of the polynomials $f_1, \cdots, f_{n-1}$.

Substituting (11) into (10′) we get

$$A_1^{l_n} f_n \equiv x_1^{s_n} \phi_n(x_1)(f_1, \cdots, f_{n-1}); \; A_1^{l_{n+1}} f_{n+1} \equiv x_1^{s_{n+1}} \phi_{n+1}(x_1)(f_1, \cdots, f_{n-1}), \tag{12}$$

where

$$\begin{aligned} \phi_n(x_1) &= A_1^{l_n - s_n} f_{n,s_n}(A_1, \cdots, A_n) + x_1 A_1^{l_n - s_n - 1} f_{n,s_n+1}(A_1, \cdots, A_n) + \cdots \\ &\quad + x_1^{l_n - s_n} f_{n,l_n}(A_1, \cdots, A_n); \\ \phi_{n+1}(x_1) &= A_1^{l_{n+1} - s_{n+1}} f_{n+1,s_{n+1}}(A_1, \cdots, A_n) \\ &\quad + x_1 A_1^{l_{n+1} - s_{n+1} - 1} f_{n+1,s_{n+1}+1}(A_1, \cdots, A_n) + \cdots \\ &\quad + x_1^{l_{n+1} - s_{n+1}} f_{n+1,l_{n+1}}(A_1, \cdots, A_n). \end{aligned}$$

Let $R=R(\phi_n, \phi_{n+1})$ be the resultant of $\phi_n(x_1)$ and $\phi_{n+1}(x_1)$. We have $R\equiv 0(\phi_n, \phi_{n+1})$ and hence, by (12), $x_1^\sigma R\equiv 0(f_1, \cdots, f_{n-1}, f_n, f_{n+1})$, for some σ. It follows, by Theorem 2, that R is a form of inertia of the polynomials f_i. From the form of the coefficients of $\phi_n(x_1)$ and $\phi_{n+1}(x_1)$ and from the fact that R is an isobaric form of weight $(l_n-s_n)(l_{n+1}-s_{n+1})$ in these coefficients, it follows that $A_1^{(l_n-s_n)(l_{n+1}-s_{n+1})}$ is a factor of R. Let $R=A_1^{(l_n-s_n)(l_{n+1}-s_{n+1})}\cdot P$. Now A_1 is independent of the coefficients of f_n and f_{n+1} and hence, by Theorem 4, is not a form of inertia of $f_1, \cdots, f_{n+1}$. Consequently *P is a form of inertia of $f_1, \cdots, f_{n+1}$*. The coefficients of $\phi_i(i=n, n+1)$ are homogeneous of degree 1 in the coefficients of f_i and homogeneous of degree l_i in $A_1, \cdots, A_n$, hence homogeneous of degree l_i in the coefficients of each of the polynomials $f_1, \cdots, f_{n-1}$. Hence R is homogeneous of degree l_n-s_n and $l_{n+1}-s_{n+1}$ in the coefficients of f_{n+1} and f_n respectively, and homogeneous of degree $l_n(l_{n+1}-s_{n+1})+l_{n+1}(l_n-s_n)$ in the coefficients of f_i, $i=1, 2, \cdots, n-1$. It follows that *P is homogeneous of degree $l_{n+1}-s_{n+1}$ and l_n-s_n in the coefficients of f_n and f_{n+1} respectively, and homogeneous of degree $l_nl_{n+1}-s_ns_{n+1}$ in the coefficients of each of the polynomials $f_1, \cdots, f_{n-1}$*.

It remains to prove that $P=D$, or, what is the same, that P is an irreducible polynomial in the coefficients of $f_1, \cdots, f_{n+1}$. We observe that P is the resultant of the following polynomials

$$\begin{aligned}\psi_n(x_1; A_1, \cdots, A_n) &= f_{n,s_n}(A_1, \cdots, A_n) + x_1 f_{n,s_n+1}(A_1, \cdots, A_n)\\ &\quad + \cdots + x_1^{l_n-s_n} f_{n,l_n}(A_1, \cdots, A_n),\\ \psi_{n+1}(x_1; A_1, \cdots, A_n) &= f_{n+1,s_{n+1}}(A_1, \cdots, A_n) + x_1 f_{n+1,s_{n+1}+1}(A_1, \cdots, A_n)\\ &\quad + \cdots + x_1^{l_{n+1}-s_{n+1}} f_{n+1,l_{n+1}}(A_1, \cdots, A_n).\end{aligned}$$

For the special polynomials $f_1=x_2$, $f_2=x_3, \cdots, f_{n-1}=x_n$, we have $A_1=1$, $A_2=\cdots=A_n=0$, and ψ_n, ψ_{n+1} become general polynomials with literal coefficients in x_1, of degree l_n-s_n and $l_{n+1}-s_{n+1}$ respectively, and their resultant is irreducible. Hence P cannot be divisible by two factors or by the square of a factor in which the coefficients of f_n or of f_{n+1} actually occur. On the other hand, for the special polynomials

$$\begin{aligned}\psi_n &= x_1^{l_n-s_n} f_{n,l_n}(A_1, \cdots, A_n),\\ \psi_{n+1} &= f_{n+1,s_{n+1}}(A_1, \cdots, A_n) + x_1^{l_{n+1}-s_{n+1}} f_{n+1,l_{n+1}}(A_1, \cdots, A_n)\end{aligned}$$

we get $P=\pm f_{n,l_n}^{l_{n+1}-s_{n+1}} f_{n+1,l_{n+1}}^{l_n-s_n}$, and hence P cannot have a factor independent of the coefficients of both f_n and f_{n+1}. Hence P is irreducible, $P=D$.

Passing to the general case where $f_1, \cdots, f_{n-1}$ are of arbitrary degrees $l_1, \cdots, l_{n-1}$, while f_n, f_{n+1} are non-homogeneous polynomials, we specialize

each polynomial f_i, $i=1, 2, \cdots, n-1$, into the product of l_i linear factors, of which s_i are linear forms: $f_i = f_{i1} f_{i2} \cdots f_{il_i}$. By the special case considered above, the irreducible form of inertia $D_{j_1 \cdots j_{n-1}}$ of the $n+1$ polynomials $f_{1j_1}, f_{2j_2}, \cdots, f_{n-1,j_{n-1}}, f_n, f_{n+1}$ $(1 \leqq j_i \leqq l_i)$ actually contains the coefficients of each factor. Hence we get $l_1 l_2 \cdots l_{n-1}$ distinct irreducible forms of inertia and their product must divide D. Now $D_{j_1 \cdots j_{n-1}}$ is of degree $l_n l_{n+1}$ in the coefficients of f_{1j_1}, if the polynomials $f_{2j_2}, \cdots, f_{n-1,j_{n-1}}$ are not all homogeneous, and is of degree $l_n l_{n+1} - s_n s_{n+1}$ in the coefficients of f_{1j_1}, if all the polynomials $f_{2j_2}, \cdots, f_{n-1,j_{n-1}}$ are homogeneous. It follows that D is of degree $\geqq l_2 \cdots l_{n+1} - s_2 \cdots s_{n+1}$ in the coefficients of f_1. Similarly D is of degree $\geqq l_1 \cdots l_{i-1} l_{i+1} \cdots l_{n+1} - s_1 \cdots s_{i-1} s_{i+1} \cdots s_{n+1}$ in the coefficients of f_i, $i=1, 2, \cdots, n-1$. $D_{j_1 \cdots j_{n-1}}$ is of degree l_{n+1} in the coefficients of f_n, if $f_{1j_1}, \cdots, f_{n-1,j_{n-1}}$ are not all homogeneous, and is of degree $l_{n+1} - s_{n+1}$ in the contrary case. Hence D is of degree $l_1 \cdots l_{n-1} l_{n+1} - s_1 \cdots s_{n-1} s_{n+1}$ in the coefficients of f_n. Similarly, D is of degree $l_1 \cdots l_n - s_1 \cdots s_n$ in the coefficients of f_{n+1}.

9. **An application to the intersection theory of algebraic hypersurfaces.** Let

$$\phi_1(x_1, \cdots, x_n, x_{n+1}) = 0,$$
$$\cdots\cdots\cdots\cdots$$
$$\phi_{n+1}(x_1, \cdots, x_n, x_{n+1}) = 0,$$

be the equations of $n+1$ hypersurfaces $F_1, \cdots, F_{n+1}$ in the $(n+1)$-dimensional projective space. Let l_i be the order of F_i. Let the origin $O(0, \cdots, 0)$ be a common point of these hypersurfaces, and let it be an s_i-fold point of F_i. We regard ϕ_i as a polynomial in $x_1, \cdots, x_n$, and we write $\phi_i = \sum a^{(i)}_{j_1 \cdots j_n} x_1^{j_1} \cdots x_n^{j_n}$, where the coefficients $a^{(i)}_{(j)}$ are polynomials in x_{n+1}. Since O is an s_i-fold point of F_i, $a^{(i)}_{j_1 \cdots j_n}$ is divisible by $x_1^{s_i - j_1 - \cdots - j_n}$, if $j_1 + \cdots + j_n \leqq s_i$. Hence, by Theorem 6, every term of the resultant $R(\phi_1, \cdots, \phi_{n+1}) = R(x_{n+1})$ is divisible by $x_{n+1}^{s_1 \cdots s_{n+1}}$. Let

$$R(x_{n+1}) = \alpha x^{s_1 \cdots s_{n+1}} + \text{terms of higher degree},$$

where α is a constant. Let g_i $(i=1, 2, \cdots, n+1)$ denote the sum of terms of lowest degree (s_i) in ϕ_i. Then we have, by Theorem 6, $\alpha = cD_0^{\sigma} R(g_1, \cdots, g_{n+1})$, where c is an integer, and $D_0 = [D]_{x_{n+1}=0}$. The homogeneous equation $g_i = 0$ represents the tangent hypercone of the hypersurface F_i at the point O. Hence $R(g_1, \cdots, g_{n+1})$ vanishes, if and only if the $n+1$ hypersurfaces F_i have a common principal tangent line at O. Assume that $R(g_1, \cdots, g_{n+1}) \neq 0$. If f_i

denotes, as in Theorem 6, the sum of terms in ϕ_i which are of degree $\geqq s_i$ in $x_1, \cdots, x_n$, then $f_i = [g_i]_{x_{n+1}=0}$ + terms of degree $> s_i$ in $x_1, \cdots, x_n$. It follows, by Theorem 3.1, that if $R(g_1, \cdots, g_{n+1}) \neq 0$, then $D_0 = 0$ implies that the hypersurfaces F_i have a common point on the hyperplane $x_{n+1} = 0$, outside the origin; and conversely, by Theorem 3.2. Assuming that the hypersurfaces F_i meet in a finite number of points, we see that if the coordinate axes are in generic position and if the hypersurfaces F_i have no principal tangent in common at the point O, then $\alpha \neq 0$. According to the usual definition of the intersection multiplicity of the hypersurfaces F_i at a common point, it follows that *the intersection multiplicity at* O *is* $\geqq s_1 \cdots s_{n+1}$ *and equals* $s_1 \cdots s_{n+1}$ *if and only if the hypersurfaces* F_i *have no common principal tangent at* O.

THE JOHNS HOPKINS UNIVERSITY,
BALTIMORE, MD.

ON THE THEOREM OF JORDAN-HÖLDER*

BY

OYSTEIN ORE

In a group we have the well known theorem on *principal series*, that any two principal maximal series have the same length and the quotient groups in the two series are isomorphic in some order. In a paper entitled *Über die von drei Moduln erzeugte Dualgruppe*, Dedekind† analyzed the axiomatic foundation of this theorem, particularly the fact that the length of two maximal principal series is the same. He showed that this theorem can be considered as a theorem on *structures* (Dualgruppen), i.e., systems with two operations called union and cross-cut. In order that the theorem shall hold in such a structure it is necessary that it satisfy a further condition which I have called the *Dedekind axiom*. Similar considerations have been made by G. Birkhoff.‡ In a recent paper§ I have shown that in Dedekind structures one can prove a general theorem corresponding to the theorem of Schreier-Zassenhaus‖ for principal series in groups. This theorem contains the analogue of the theorem of Jordan-Hölder for Dedekind structures and yields also the fact that the quotients are isomorphic in some order.

All these investigations apply only to Dedekind structures, and give the analogues to the theorems on principal series, i.e., series of sub-groups where each group is a normal sub-group of the full group. They do *not* apply to *composition series* where one only supposes that each term is normal under the preceding. In this paper we shall investigate the possibility of deriving a theory applicable to arbitrary structures and giving an analogue to the theorem of Jordan-Hölder for composition series. The first step is to examine the validity of the analogue to the second theorem of isomorphism (Theorem 1). Next we have to introduce some notion of normality and normal elements. It turns out that two suitable types of normality, α- and β-normality may be

---

* Presented to the Society, September 5, 1936; received by the editors May 6, 1936.

† Mathematische Annalen, vol. 53 (1900), pp. 371–403; Werke, vol. 2, pp. 236–271.

‡ Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464; vol. 30 (1934), pp. 115–122, p. 200.

§ Oystein Ore, *On the foundation of abstract algebra*, I, Annals of Mathematics, vol. 36 (1935), pp. 406–437. For further elucidation on the concepts used in the following I shall have to refer to this paper.

‖ O. Schreier, *Über den Jordan-Hölderschen Satz*, Abhandlungen aus dem Mathematischen Seminar, Hamburg, vol. 6 (1928), pp. 300–302.

H. Zassenhaus, *Zum Satz von Jordan-Hölder-Schreier*, Abhandlungen aus dem Mathematischen Seminar, Hamburg, vol. 10 (1934), pp. 106–108.

defined, each corresponding to some particular property of the decomposition theorem. For normal sub-groups both properties are always satisfied. The main theorem is then Theorem 7, which gives the analogue of the Schreier-Zassenhaus theorem for composition series. In the last part I discuss the difference between the theorems in structures and the corresponding theorems for groups.

1. Structures and quotients. We shall in the following consider an arbitrary *structure* Σ, i.e., an algebraic system consisting of elements $A, B, \cdots$ with an inclusion relation $A > B$ holding for certain pairs of elements. Furthermore, we suppose that to any two elements A and B there exists a *union* $[A, B]$ which is a minimal element of Σ containing A and B and a *cross-cut* (A, B) which is a maximal element contained in A and B. For these symbols we have the ordinary axioms

$$\begin{aligned} &(A, B) = (B, A), && [A, B] = [B, A], \\ &(A, A) = A, && [A, A] = A, \\ &(A, (B, C)) = ((A, B), C), && [A, [B, C]] = [[A, B], C], \\ &[A, (A, B)] = A, && (A, [A, B]) = A. \end{aligned}$$

We shall say Σ has a *unit element* E_0 and an *all-element* O_0 if these elements satisfy the relations

$$[A, O_0] = O_0, \qquad (A, E_0) = E_0$$

for all A in Σ.

The structures Σ and Σ' shall be said to be *structure isomorphic* if there exists a one-to-one correspondence $A \rightleftarrows A'$ between the elements of the two structures such that if

$$A \to A', \qquad B \to B'$$

then

$$(A, B) \to (A', B'), \qquad [A, B] \to [A', B'].$$

To any two elements $A > B$ it is convenient to associate a symbol, the *quotient* $\mathfrak{A} = A/B$. These quotients may themselves be made into a structure by defining that for

$$\mathfrak{A}_1 = A_1/B_1, \qquad \mathfrak{A}_2 = A_2/B_2$$

we shall have

$$[\mathfrak{A}_1, \mathfrak{A}_2] = [A_1, A_2]/[B_1, B_2], \qquad (\mathfrak{A}_1, \mathfrak{A}_2) = (A_1, A_2)/(B_1, B_2).$$

We shall usually apply these operations only in the case where $\mathfrak{A}_1$ and $\mathfrak{A}_2$ have the same denominator or the same numerator. Furthermore if

$$\mathfrak{A} = A/B, \qquad \mathfrak{L} = B/C$$

we define the *product*

$$\mathfrak{A} \times \mathfrak{L} = A/B \times B/C = A/C.$$

The existence of a *chain*

$$A_1 \geqq A_2 \geqq \cdots \geqq A_{n+1}$$

may then also be expressed by saying that the quotient

$$\mathfrak{A} = A_1/A_{n+1}$$

has a factorization

$$\mathfrak{A} = \mathfrak{A}_1 \times \mathfrak{A}_2 \times \cdots \times \mathfrak{A}_n, \qquad \mathfrak{A}_i = A_i/A_{i+1}.$$

Furthermore we associate with each quotient $\mathfrak{A} = A/B$ a *quotient structure* consisting of all elements S in Σ satisfying the condition

$$A \geqq S \geqq B.$$

We have formerly studied in detail the so-called *Dedekind structures*, i.e., structures satisfying the

Dedekind Axiom. *For any three elements A, B, C in Σ such that $C > A$ we have*

$$(C, [A, B]) = [A, (C, B)].$$

In the following we shall not suppose that the Dedekind axiom is satisfied. In Dedekind structures important considerations were based upon the notion of *transformation*. A large part of this theory may also be developed in structures not satisfying the Dedekind axiom. We define for any two quotients with the same denominator

$$\mathfrak{A} = A/B, \qquad \mathfrak{T} = C/B \tag{1}$$

the (right-hand) *transform of* $\mathfrak{A}$ *by* $\mathfrak{T}$ to be the quotient

$$\mathfrak{A}' = \mathfrak{T}\mathfrak{A}\mathfrak{T}^{-1} = [\mathfrak{A}, \mathfrak{T}] \times \mathfrak{T}^{-1} = [A, C]/C. \tag{2}$$

We mention without proof that most of the theorems established for transformations in Dedekind structures will also hold in general structures.

As before we shall call (2) an *extension* of $\mathfrak{A}$ by $\mathfrak{T}$ if $\mathfrak{A}$ and $\mathfrak{T}$ in (1) are relatively prime, i.e., if in (1) we have $(A, C) = B$. Conversely, we shall call $\mathfrak{A}$ in (2) a *contraction* of $\mathfrak{A}'$. A series of extensions and contractions shall be

called a *similarity transformation* of $\mathfrak{A}$, and $\mathfrak{A}$ and $\mathfrak{B}$ are said to be *similar* when one can be obtained from the other by a similarity transformation.

2. **The second law of isomorphism.** A fundamental result for Dedekind structures was the result that similar quotients were associated with isomorphic quotient structures. The proof for this fact was based upon the fundamental theorem that the two quotient structures

$$\mathfrak{A} = [A, B]/A, \qquad \mathfrak{B} = B/(A, B) \tag{3}$$

were isomorphic. This is the analogue of the so-called second law of isomorphism for groups and ideals.

Let us now determine some condition for the two quotients (3) to be isomorphic even when the Dedekind axiom is not satisfied in Σ. We denote by $\overline{A}$ and $\overline{B}$ arbitrary elements such that

$$[A, B] \geqq \overline{A} \geqq A, \qquad B \geqq \overline{B} \geqq (A, B). \tag{4}$$

One can then easily obtain a correspondence between the two quotients $\mathfrak{A}$ and $\mathfrak{B}$ by putting

$$\overline{A} \to (\overline{A}, B), \qquad \overline{B} \to [A, \overline{B}]. \tag{5}$$

We shall call (5) the *regular correspondence* between $\mathfrak{A}$ and $\mathfrak{B}$. We can then prove:

THEOREM 1. *The necessary and sufficient condition for the regular correspondence to establish a structure isomorphism between the quotients* (3) *is that for every $\overline{A}$ and $\overline{B}$ defined by* (4) *we have*

$$\overline{A} = [A, (B, \overline{A})], \qquad \overline{B} = (B, [A, \overline{B}]). \tag{6}$$

The conditions (6) are obviously necessary and sufficient in order that the regular correspondence be a one-to-one correspondence, one correspondence (5) being the inverse of the other. To prove that it also establishes an isomorphism let

$$\overline{B}_1 \to \overline{A}_1 = [A, \overline{B}_1], \qquad \overline{B}_2 \to \overline{A}_2 = [A, \overline{B}_2].$$

Then obviously

$$[\overline{B}_1, \overline{B}_2] \to [A, [\overline{B}_1, \overline{B}_2]] = [\overline{A}_1, \overline{A}_2].$$

In the same way

$$(\overline{B}_1, \overline{B}_2) \to [A, (\overline{B}_1, \overline{B}_2)] = [A, (\overline{A}_1, \overline{A}_2, B)] = (\overline{A}_1, \overline{A}_2).$$

Let us observe that this proof also simplifies the demonstration of Theorem 1 in the case of Dedekind structures.

3. **Types of normality.** In order to obtain for arbitrary structures an analogue to the theorem of Jordan-Hölder it is necessary to introduce some notion of *normality* and *normal element.* It is of interest that this may be done in several ways and that each such condition of normality has some particular meaning for the theorem of Jordan-Hölder. We shall begin by defining:

α. *An element A_0 contained in M shall be said to be α-normal in M if it satisfies the condition:*

For any $B \geqq C$ contained in M we have

$$(B, [A_0, C]) = [C, (B, A_0)]. \tag{7}$$

This condition for α-normality may be formulated in various equivalent ways obtainable from (7) by a suitable choice of B and C. We mention only:

α'. *For every B and C contained in M we have*

$$\begin{aligned} ([B, C], [A_0, C]) &= [C, (A_0, [B, C])], \\ (B, [A_0, (B, C)]) &= [(B, C), (A_0, B)], \\ ([B, C], [A_0, (B, C)]) &= [(B, C), (A_0, [B, C])]. \end{aligned} \tag{8}$$

Furthermore:

α''. *If $B \geqq C$ are contained in M and*

$$[A_0, C] = [A_0, B], \qquad (A_0, C) = (A_0, B),$$

then we can conclude $B = C$.

We may also define α-normality of A_0 with respect to elements not containing A_0 in the following manner:

An element A_0 is said to be α-normal with respect to B if A_0 is α-normal in $[B, A_0]$.

Obviously if A_0 is α-normal with respect to B it is α-normal with respect to any element contained in B. We can now prove

THEOREM 2. *If A_0 is α-normal with respect to B, i.e., A_0 is α-normal in $[A_0, B]$, then (A_0, B) is α-normal in B.*

Let B_1 and B_2 be any elements such that

$$B \geqq B_1 \geqq B_2.$$

We find then

$$(B_1, [B_2, (B, A_0)]) = (B_1, [B_2, A_0]) = [B_2, (B_1, A_0)] = [B_2, (B_1, (B, A_0))]$$

and our theorem is proved.

A second type of normality may be introduced as follows:

β. *An element A_0 contained in M shall be said to be β-normal in M when it satisfies the condition*:

For any B and C contained in M such that $B \geqq A_0$ we shall have

$$(B, [A_0, C]) = [A_0, (B, C)]. \tag{9}$$

This condition for β-normality may again be stated in other equivalent forms, for instance:

β'. *For any B and C contained in M we have*

$$([B, A_0], [C, A_0]) = [A_0, (B, [A_0, C])] = [A_0, (C, [A_0, B])].$$

Also here we can define normality of A_0 with respect to an arbitrary element.

THEOREM 3. *If A_0 and A_1 are both β-normal in M then $[A_0, A_1]$ has the same property.*

If, namely, $B \geqq [A_0, A_1]$ then

$$(B, [A_0, A_1, C]) = [A_0, (B, [A_1, C])] = [A_0, A_1, (B, C)].$$

4. **Semi-normality.** We shall now join the two notions of α-normality and β-normality and define:

Semi-normality. An element A_0 is semi-normal in M if it is both α-normal and β-normal in M.

We say further that A_0 is semi-normal with respect to B if it is semi-normal in $[B, A_0]$.

THEOREM 4. *The necessary and sufficient condition for A_0 to be semi-normal in M is that for any B and C in M we have*

$$[(A_0, [B, C]), (C, [B, A_0])] = ([A_0, B], [A_0, C], [B, C]). \tag{10}$$

We obtain namely by supposing $B > A_0$

$$(B, [A_0, C]) = [A_0, (B, C)]$$

and by supposing $B > C$

$$(B, [A_0, C]) = [C, (B, A_0)],$$

and conversely these relations suffice to derive (10).

The principal theorem on semi-normality is:

THEOREM 5. *When A_0 is semi-normal with respect to B; i.e., A_0 is semi-normal in $[B, A_0]$, then (B, A_0) is semi-normal in B and the two quotients*

$$\mathfrak{A} = [A_0, B]/A_0, \qquad \mathfrak{B} = B/(A_0, B), \tag{11}$$

are structure isomorphic.

It follows from Theorem 2 that (A_0, B) is α-normal in B. The β-normality implies that for any B_1 and B_2 in B such that $B_1 \geqq (A_0, B)$ we shall have

$$(B_1, [(A_0, B), B_2]) = [(A_0, B), (B_1, B_2)]. \tag{12}$$

To prove this relation we observe that on account of the α-normality of A_0 we have

$$B_1 = (B, [A_0, B_1])$$

and hence

$$[(A_0, B), (B_1, B_2)] = [(A_0, B), (B_2, [A_0, B_1])] = (B, [A_0, (B_2, [A_0, B_1])]).$$

The β-normality shows that this last expression is identical with

$$(B, [A_0, B_1], [A_0, B_2]) = (B_1, [A_0, B_2]) = (B_1, [(A_0, B), B_2])$$

and (12) is proved.

The structure isomorphism between the two quotients (11) follows from Theorem 1. For any A_1 such that

$$[A_0, B] \geqq A_1 \geqq A_0$$

we have, namely,

$$[A_0, (A_1, B)] = (A_1, [A_0, B]) = A_1,$$

and similarly for any B_1 such that

$$B \geqq B_1 \geqq (A_0, B)$$

we have

$$(B, [A_0, B_1]) = [B_1, (B_1, A_0)] = B_1.$$

5. **The theorem of Schreier-Zassenhaus.** We shall now consider the possibility of extending the theorem of Jordan-Hölder, or rather its generalization by Schreier-Zassenhaus, to an arbitrary structure Σ. Let us suppose that we have two descending chains between two elements A and B,

$$\begin{aligned} &A > B_1 > \cdots > B_{r-1} > B, \\ &A > C_1 > \cdots > C_{s-1} > B, \end{aligned} \tag{13}$$

or in the terminology of §1, that the quotient $\mathfrak{A} = A/B$ has two product representations

$$\mathfrak{A} = \mathfrak{B}_1 \times \mathfrak{B}_2 \times \cdots \times \mathfrak{B}_r = \mathfrak{C}_1 \times \mathfrak{C}_2 \times \cdots \times \mathfrak{C}_s \tag{14}$$

where

$$\mathfrak{B}_i = B_{i-1}/B_i, \qquad \mathfrak{C}_j = C_{j-1}/C_j.$$

We can now prove

THEOREM 6. *Let there exist two sequences* (13) *between A and B where each term is α-normal under the preceding. In the corresponding factorizations* (14) *of the quotient $\mathfrak{A}=A/B$ it is then possible to decompose the factors further*

$$\mathfrak{B}_i = \mathfrak{B}_{i,1} \times \mathfrak{B}_{i,2} \times \cdots \times \mathfrak{B}_{i,s}, \qquad \mathfrak{C}_j = \mathfrak{C}_{j,1} \times \mathfrak{C}_{j,2} \times \cdots \times \mathfrak{C}_{j,r} \tag{15}$$

such that the two factors $\mathfrak{B}_{i,j}$ and $\mathfrak{C}_{j,i}$ correspond in the manner that they may both be obtained by extension from the same quotient $\mathfrak{R}_{i,j}$.

We put

$$\begin{aligned} \mathfrak{B}_{i,j} &= [B_i, (B_{i-1}, C_{j-1})]/[B_i, (B_{i-1}, C_j)], \\ \mathfrak{C}_{j,i} &= [C_i, (C_{i-1}, B_{j-1})]/[C_i, (C_{i-1}, B_j)] \end{aligned} \tag{16}$$

and write

$$L = [B_i, (B_{i-1}, C_j)], \qquad M = (B_{i-1}, C_{j-1}).$$

We then find

$$[L, M] = [B_i, (B_{i-1}, C_{j-1})], \qquad \mathfrak{B}_{i,j} = [L, M]/L$$

and on account of the α-normality of B_i in B_{i-1} we have

$$(L, M) = (B_{i-1}, C_{j-1}, [B_i, (B_{i-1}, C_j)]) = [(B_{i-1}, C_j), (B_i, C_{j-1})].$$

This shows that $\mathfrak{B}_{i,j}$ may be obtained by extension from

$$\mathfrak{R}_{i,j} = M/(L, M) = (B_{i-1}, C_{j-1})/[(B_{i-1}, C_j), (B_i, C_{j-1})],$$

and since similar considerations show that $\mathfrak{C}_{j,i}$ may be obtained by extension from $\mathfrak{R}_{i,j}$, our theorem is proved.

We also mention without proof that, when α-normality is assumed in the sequences (13), repeated applications of the decompositions (16) yield no new factorizations.

Under the assumption of semi-normality we can prove the more exact theorem:

THEOREM 7. *Let there exist two chains* (13) *between A and B such that each term is semi-normal under the preceding. The corresponding factorizations* (14) *of $\mathfrak{A}=A/B$ may then be factored further into* (15) *such that the new factors $\mathfrak{B}_{i,j}$ and $\mathfrak{C}_{j,i}$ have isomorphic quotient structures.*

Since we have

$$\mathfrak{B}_{i,j} = [L, M]/L, \qquad \mathfrak{R}_{i,j} = M/(M, L),$$

it follows from Theorem 5 that $\mathfrak{B}_{i,j}$ and $\mathfrak{R}_{i,j}$ are structure isomorphic. Since the same holds for $\mathfrak{R}_{i,j}$ and $\mathfrak{C}_{j,i}$ our theorem is proved. We can obtain the explicit correspondence between $\mathfrak{B}_{i,j}$ and $\mathfrak{C}_{j,i}$ in the following way: Let $\overline{B}$ and $\overline{C}$ be arbitrary elements such that

$$[B_i, (B_{i-1}, C_{j-1})] \geqq \overline{B} \geqq [B_i, (B_{i-1}, C_j)],$$
$$[C_j, (C_{j-1}, B_{i-1})] \geqq \overline{C} \geqq [C_j, (C_{j-1}, B_i)].$$

We obtain the correspondence by putting

$$\overline{B} \to [C_j, (C_{j-1}, \overline{B})], \qquad \overline{C} \to [B_i, (B_{i-1}, \overline{C})].$$

6. **Comparison with normal sub-groups.** In the preceding we have tried to derive results similar to those known in the case of normal sub-groups. The main theorem, 7, is almost identical with the theorem of Schreier-Zassenhaus for groups.

Let us now take the opposite view of the matter and consider the difference of our results from the group theorems. To any two semi-normal chains (13) in a structure we have constructed new chains by intercalation such that the quotients in the two new chains are structure isomorphic in pairs. Due to our weak condition of semi-normality we can not, however, prove that in the new decompositions each term is semi-normal under the preceding.

In the case of groups this deficiency is easily remedied. To show that the group

$$D_{i,j-1} = [B_i, (B_{i-1}, C_{j-1})]$$

contains the normal sub-group

$$D_{i,j} = [B_i, (B_{i-1}, C_j)]$$

when B_i is normal in B_{i-1} and C_j normal in C_{j-1}, it is only necessary to show that B_i and (B_{i-1}, C_j) are transformed into themselves by transformation with elements in B_i and (B_{i-1}, C_j). This follows directly from the definition of normality in groups.

On account of this difference Theorem 7 may not be specialized into the analogue of the theorem of Jordan-Hölder by supposing that in the chains (13) each term is maximal semi-normal under the preceding. Let us try to determine however some conditions under which the theorem of Jordan-Hölder is valid for structures.

In order to do this, let us consider the structure formed by *all* sub-groups of a given group $\mathfrak{G}$. We then have:

I. The set of normal sub-groups of a given sub-group M forms a Dedekind structure.

II. For a given sub-group A those sub-groups M in which A is normal form a structure.

III. Let A be normal in $[A, B]$. The structure isomorphism

$$[A, B]/A \cong B/(A, B)$$

is also a one-to-one correspondence between the normal sub-groups of $[A, B]$ containing A and the normal sub-groups of B containing (A, B).

None of these properties is ordinarily satisfied for semi-normal elements in structures. For instance, the set of semi-normal elements in M usually does not even form a structure. With regard to III one can prove that every semi-normal element in $[A, B]/A$ corresponds to a semi-normal element in $B/(A, B)$ but not conversely.

For the proof of an analogue to the theorem of Jordan-Hölder one needs III and a part of I, namely, that the union of two semi-normal elements in M is again semi-normal. If one then has two series (13) in which each term is maximal semi-normal in the preceding, the ordinary inductional proof carries through without difficulty. It is an interesting fact that one does not need all the properties of the normal sub-groups and I shall use this fact to prove in a following paper a new theorem of Jordan-Hölder which is valid also for certain classes of non-normal sub-groups.

YALE UNIVERSITY,
NEW HAVEN, CONN.

LINEAR DIVISIBILITY SEQUENCES*

BY

MORGAN WARD

I. Introduction

1. A sequence of rational integers

$$(u): \qquad u_0, u_1, \cdots, u_n, \cdots$$

is called a *divisibility sequence* if u_n divides u_m whenever n divides m. (u) is *linear*† if it satisfies a linear difference equation with integral coefficients and *normal* if $u_0=0$, $u_1=1$. Marshall Hall has shown that a linear divisibility sequence is usually normal [2]. If

$$(1.1) \qquad f(x) = x^k - c_1x^{k-1} - \cdots - c_k, \qquad c_1, \cdots, c_k \text{ integers},$$

is the polynomial associated with the difference equation of lowest order which (u) satisfies, (u) is said to be of *order* k and to *belong to* its *characteristic polynomial* $f(x)$.

An integer dividing every term of (u) beyond a certain point is called a *null divisor* of (u) [3]. If (u) has no null divisors save ± 1, it is said to be *primary*.

If u_s is any fixed non-vanishing term of (u), the sequence

$$u_0/u_s, u_s/u_s, u_{2s}/u_s, \cdots, u_{ns}/u_s, \cdots$$

is called a *subsequence* of (u). The various subsequences of (u) are themselves normal linear divisibility sequences of order $\leqq k$.

2. The object of this paper is to prove the following results:

Let the characteristic polynomial of the linear divisibility sequence (u) *have no repeated roots, and let its coefficients be relatively prime. Then*:

I. *If* (u) *is primary and if* q *is any large prime number,*

$$(2.1) \qquad u_q^\sigma \equiv 1 \pmod q,$$

where σ *is the least common multiple of* $1, 2, 3, \cdots, k$.

II. *If* (u) *is not primary it always contains an infinity of subsequences which are primary. Furthermore the characteristic polynomials of such subsequences satisfy the hypotheses imposed above upon the polynomial* (1.1).

* Presented to the Society, June 18, 1936; received by the editors May 5, 1936.

† T. A. Pierce appears to have been the first to discuss sequences of order greater than two [1]. (Numbers in square brackets refer to the bibliography at the end of the paper.)

III. *There exists a rational number*

$$B = B(u) = B(u_0, u_1, \cdots, u_{k-1}; c_1, \cdots, c_k) = \frac{P}{Q}, \qquad (P, Q) = 1$$

such that

(i) *if p is a prime number dividing neither the numerator* P *nor the denominator* Q *of B, then the rank of apparition* of p in the sequence (u) is the restricted period* of (u) modulo p;*

(ii) *the prime factors of the denominator of B all divide the discriminant of the polynomial to which (u) belongs;*

(iii) *the numerator of B can never vanish if the galois group of $f(x)$ is alternating or symmetric.*†

II. Proof of first result

3. Given any modulus m, the least period of (u) modulo m is called its characteristic number and the number of non-periodic terms in (u) modulo m its numeric. The reader will be assumed to be familiar with my previous paper in these Transactions [4] (referred to hereafter as T) devoted to the determination of these numbers.

Henceforth let (u) be a normal linear divisibility sequence of order k, and let D denote the discriminant of its characteristic polynomial. We assume:

$$D \neq 0. \tag{3.1}$$

Lemma 3.1 [4]. *If*‡ $(q, D) = 1$, *q a prime, and if σ is the least common multiple of* $2, 3, \cdots, k$, *then (u) admits the period $q^\sigma - 1$ modulo q.*

Theorem 3.1. *If (u) is a linear divisibility sequence of order k and q a prime such that $u_q \equiv 0$ (mod q), then either q divides D or q divides c_k.*

Assume that‡ $q \mid u_q$, q a prime. The assumption $(q, c_k) = (q, D) = 1$ then yields a contradiction. For if $(q, c_k) = 1$, (u) is purely periodic modulo q [5]. And if $(q, D) = 1$, (u) admits the period $q^\sigma - 1$ modulo q. Determine positive integers x and y such that $xq = y(q^\sigma - 1) + 1$. Then $u_{xq} \equiv u_1 \equiv 1$ (mod q). But $q \mid u_q$ and $u_q \mid u_{xq}$.

The following lemma is a direct consequence of Theorem 3.1.

---

* The rank of apparition of p is the index ρ of the first term of (u) excluding u_0 which divides: $u_\tau \equiv 0$ (mod p); $u_n \not\equiv 0$ (mod p), $0 < n < \rho$. The restricted period [5] of (u) modulo is the least positive integer τ such that $u_{n+\tau} \equiv cu_n$ (mod p), $n = 0, 1, 2, \cdots$, c an integer. ρ always divides τ [2].

† It is unknown whether divisibility sequences exist whose characteristic polynomial is restricted as in (iii). No such sequences exist when $k = 3$ [2].

‡ If $a, b, c, \cdots$ are rational integers, we write as usual $(a, b, c, \cdots)$ for the greatest common divisor of $a, b, c, \cdots$, and $a \mid b$ for a divides b.

LEMMA 3.2. *There exists a rational integer q_0 such that*

$$u_q \not\equiv 0 \pmod{q}, \quad q \text{ a prime} \geqq q_0. \tag{3.2}$$

LEMMA 3.3 [4]. *For any prime p, $p^k(p^\sigma - 1)$ is a period of (u) modulo p.*

LEMMA 3.4 [4]. *For any prime p, the numeric of (u) modulo p is less than or equal to k.*

THEOREM 3.2. *If p is a prime dividing a term u_q of the divisibility sequence (u) with a sufficiently large prime index q, then either*

$$p^\sigma \equiv 1 \pmod{q} \tag{3.3}$$

or else (u) is a null sequence modulo p.

Choose a prime $q > k$ and q_0 of (3.2), and assume that $u_q \equiv 0 \pmod{p}$, p a prime. By (3.2), $p \neq q$. Hence if $(p^\sigma - 1, q) = 1$, for each positive integer r there exist positive integers x, y, z such that

$$xq + yp^k(p^\sigma - 1) = r + zp^k(p^\sigma - 1). \tag{3.4}$$

By Lemma 3.3, $p^k(p^\sigma - 1)$ is a period of (u) modulo p. Therefore if $r > k$, (3.4) and Lemma 3.4 give $u_{xq} \equiv u_r \pmod{p}$. Since $p \mid u_q$ and $u_q \mid u_{xq}$, $u_r \equiv 0 \pmod{p}$ so that (u) is a null sequence modulo p.

THEOREM 3.3. *If the linear divisibility sequence (u) is primary, and if k is its order and σ the least common multiple of the numbers $2, 3, \cdots, k$, then for all sufficiently large prime indices q we have*

$$u_q^\sigma \equiv 1 \pmod{q}. \tag{2.1}$$

Choose the prime $q > k$ and q_0 of (3.2), and let the factorization of u_q be $u_q = p_1^{a_1} p_2^{a_2} \cdots p_t^{a_t}$. Since (u) is assumed primary none of the primes p_i are null divisors. Therefore Theorem 3.2, $p_i^\sigma \equiv 1 \pmod{q}$, so that

$$p_i^{\sigma a_i} \equiv 1 \pmod{q}, \qquad (i = 1, 2, \cdots, t).$$

On multiplying these t congruences together, we obtain (2.1), and our first result is proved.

III. PROOF OF SECOND RESULT

4. We assume that (u) is a normal linear divisibility sequence for which

$$(c_1, c_2, \cdots, c_k) = 1. \tag{4.1}$$

A *proper* null divisor of a linear sequence is one which divides neither its initial terms nor the coefficients of its recursion. Any other null divisor is called *trivial*. (u) obviously has no trivial null divisors.

THEOREM 4.1. *No subsequence of* (u) *has trivial null divisors.*

LEMMA 4.1 (Schatanovskis Principle) [6, 7, 8]. *If* $\Phi(x_1, x_2, \cdots, x_k)$ *is an integral symmetric function of the arguments* $x_1, \cdots, x_k$ *with integral coefficients, and if for a natural number* m

$$f(x) \equiv (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_k) \equiv (x-\gamma_1)(x-\gamma_2)\cdots(x-\gamma_k) \pmod{m},$$

where $f(x)$ *is a polynomial with integral coefficients, then*

$$\Phi(\alpha_1, \alpha_2, \cdots, \alpha_k) \equiv \Phi(\gamma_1, \gamma_2, \cdots, \gamma_k) \pmod{m}.$$

LEMMA 4.2. *Let*

$$f^{(s)}(x) = x^k - d_1x^{k-1} - \cdots - d_k$$

be the polynomial whose roots are the sth powers of the roots of $f(x)$, *and* p *a prime number. Then if* t *is any positive integer* $\leqq k$, (A) $p \mid (c_k, c_{k-1}, \cdots, c_{k-t+1})$ *when and only when* (B) $p \mid (d_k, d_{k-1}, \cdots, d_{k-t+1})$.

Assume that (A) holds. Then

$$f(x) \equiv g(x) = x^{k-t}(x^t - c_1x^{t-1} - \cdots - c_{k-t}) \pmod{p}.$$

Let the k roots of $g(x)=0$ be $\gamma_1, \gamma_2, \cdots, \gamma_t; \gamma_{t+1}=\gamma_{t+2}=\cdots=\gamma_k=0$. If the roots of $f(x)=0$ are $\alpha_1, \alpha_2, \cdots, \alpha_k$, then $d_i=\Phi(\alpha_1, \alpha_2, \cdots, \alpha_k)$, where Φ is a symmetric polynomial in its arguments with rational integral coefficients. Hence by the preceding lemma

$$d_i \equiv \Phi(\gamma_1, \gamma_2, \cdots, \gamma_k) \pmod{p}.$$

But if $g^{(s)}(x)=x^k-e_1x^{k-1}-\cdots-e_k$ is the equation whose roots are the sth powers of the roots of $g(x)=0$, then

$$e_i = \Phi(\gamma_1, \gamma_2, \cdots, \gamma_k) = \Sigma\gamma_1^s\gamma_2^s\cdots\gamma_i^s = 0 \text{ if } i > k-t.$$

Hence $d_i\equiv 0 \pmod{p}$ if $i>k-t$, so that (B) follows.

To prove the converse, it suffices to show that (A) and $c_{k-t}\not\equiv 0 \pmod{p}$ imply that $d_{k-t}\not\equiv 0 \pmod{p}$. But by what precedes,

$$d_{k-t} \equiv \Sigma(\gamma_1\gamma_2\cdots\gamma_t)^s \equiv (\gamma_1\gamma_2\cdots\gamma_t)^s \equiv c_{k-t}^s \pmod{p}.$$

Proof of Theorem 4.1. With the notation of Lemma 4.2, any subsequence (v): $v_n=u_{ns}/u_s$ of (u) is normal, so that the only possible trivial null divisors of (v) are common divisors of $d_1, d_2, \cdots, d_k$. On taking $t=k$ in Lemma 4.2, we see that if $(c_1, c_2, \cdots, c_k)=1$ then $(d_1, d_2, \cdots, d_k)=1$.

5. We begin our discussion of the proper null divisors of (u) by restating some properties of linear sequences used in T. Let

$$f_0(x) = 0, \quad f_r(x) = x^r - c_1x^{r-1} - \cdots - c_r, \qquad (r = 1, 2, \cdots, k).$$

The polynomial

$$(5.1)\qquad u(x) = u_0 f_{k-1}(x) + u_1 f_{k-2}(x) + \cdots + u_{k-1} f_0(x)$$

is called the *generator* of the sequence (u).* If furthermore

$$(5.2)\qquad \Delta(u) = \begin{vmatrix} u_0, & u_1, \cdots, & u_{k-1} \\ u_1, & u_2, \cdots, & u_k \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ u_{k-1}, & u_k, \cdots, & u_{2k-2} \end{vmatrix},$$

then

$$(5.3)\qquad \Delta(u) = (-1)^{k(k-1)/2} \operatorname{Res}\{u(x), f(x)\} = \beta_1\beta_2 \cdots \beta_k D,$$

where $u_n = \beta_1\alpha_1^n + \cdots + \beta_k\alpha_k^n$ and $\alpha_1, \cdots, \alpha_k$ are the roots of $f(x)$. Since (u) is of order k and $D \neq 0$, $\Delta(u) \neq 0$.

Consider next the $k+1$ greatest common divisors

$$\begin{aligned} e_0 &= (u_0, u_1, u_2, \cdots, u_{k-1}) \\ e_1 &= (c_k, u_1, u_2, \cdots, u_{k-1}) \\ e_2 &= (c_k, c_{k-1}, u_2, \cdots, u_{k-1}) \\ &\cdots\cdots\cdots\cdots \\ e_{k-1} &= (c_k, c_{k-1}, c_{k-2}, \cdots, u_{k-1}) \\ e_k &= (c_k, c_{k-1}, c_{k-2}, \cdots, c_1). \end{aligned}$$

Then

$$e_0 = e_1 = e_k = 1.$$

The following lemma easily follows from formula (5.1) and the results of part IV of T.

LEMMA 5.1. *Necessary and sufficient conditions that a linear sequence of order k be primary are that the $k+1$ greatest common divisors e_i be all equal to unity.*

THEOREM 5.1. *If the prime p is a null divisor of the normal linear divisibility sequence (u), then p divides both $\Delta(u)$ and the discriminant D of the characteristic polynomial $f(x)$ of (u).*

It is easily shown that every such p must divide one or the other of the numbers e_i. Since $e_k = 1$, $p \mid u_{k-1}$. Hence $p \mid u_k$, $p \mid u_{k+1}, \cdots$ by Lemma 3.4.

* We have the identity $u(x)/f(x) = \sum_0^\infty u_n/x^{n+1}$ for $|x|$ large. See T, p. 606, and [3].

Hence $p|\Delta(u)$ by formula (5.2). Since $e_0=e_1=1$, $p|c_k$ and $p|c_{k-1}$. Hence $x=0$ is a multiple root of the congruence $f(x)\equiv 0 \pmod{p}$ and $p|D$.

As a corollary, we have

LEMMA 5.2. *A sufficient condition that the divisibility sequence (u) be primary is that D and $\Delta(u)$ be co-prime.*

If p is a prime proper null divisor of (u), the exponent of the highest power of p which is a null divisor of (u) is called the *index* of p in (u) [3].

LEMMA 5.3 [3]. *Let (u) be a linear sequence for which* (4.1) *holds. Then the index of any prime null divisor p is $\leqq r$, where p^r is the highest power of p dividing $\Delta(u)$.*

THEOREM 5.2. *A subsequence of a normal linear divisibility sequence can have no prime null divisor which is not a possible null divisor of (u) itself.*

Every prime null divisor of (u) must divide c_k in (1.1) [5]. Let (v) be any subsequence of (u). By Theorem 4.1, (v) can have only proper null divisors. Hence any prime null divisor of (v) must divide the constant term d_k of the polynomial to which (v) belongs. But obviously d_k divides some power of c_k.

6. Let $f^{(s)}(x)=(x-\alpha_1^s)\cdots(x-\alpha_k^s)$ be the polynomial whose roots are the sth powers of the roots of $f(x)$, and let $D^{(s)}$ be its discriminant. $D^{(s)}/D$ is clearly an integer.

THEOREM 6.1. *The integer s may be chosen in an infinite number of ways so that $D^{(s)}/D$ is prime to D.*

Let p be any prime factor of D, $\mathfrak{F}$ the Galois field of $f(x)$, and $\mathfrak{p}$ a prime ideal factor of p in $\mathfrak{F}$. Then since $D^{1/2}=\prod_{i<j}(\alpha_i-\alpha_j)$, $p|D$ only when $\alpha_i-\alpha_j\equiv 0 \pmod{\mathfrak{p}}$ for some values of the subscripts i and j.

Now

$$\left(\frac{D^{(s)}}{D}\right)^{1/2}=\prod_{i<j}\frac{\alpha_i^s-\alpha_j^s}{\alpha_i-\alpha_j} \quad\text{and}\quad \frac{\alpha_i^s-\alpha_j^s}{\alpha_i-\alpha_j}\equiv s \pmod{[\alpha_i-\alpha_j]}.^*$$

Hence if $\alpha_i-\alpha_j\equiv 0 \pmod{\mathfrak{p}}$, then $\alpha_i^s-\alpha_j^s/(\alpha_i-\alpha_j)\equiv 0 \pmod{\mathfrak{p}}$ if and only if $s\equiv 0 \pmod{\mathfrak{p}}$; that is, if and only if $s\equiv 0 \pmod{p}$. Choose s prime to D. Then if $D^{(s)}/D$ and D have a common factor, and hence a common prime factor p, we must have for some k and l

$$\text{(6.1)}\qquad \alpha_k^s\equiv\alpha_l^s \pmod{\mathfrak{p}}, \qquad\qquad \text{(6.11)}\qquad \alpha_k\not\equiv\alpha_l \pmod{\mathfrak{p}},$$

where $\mathfrak{p}|p$. If both (6.1) and (6.11) hold, then

* The square bracket denotes a principal ideal.

(6.2) $$(\alpha_k, \mathfrak{p}) = (\alpha_l, \mathfrak{p}) = (\alpha_k - \alpha_l, \mathfrak{p}) = \mathfrak{o},$$

where $\mathfrak{o}$ as usual is the unit ideal of $\mathfrak{F}$.

Now for each pair of distinct roots α_i, α_j of $f(x)$ for which $(\alpha_i, \mathfrak{p}) = (\alpha_j, \mathfrak{p}) = (\alpha_i - \alpha_j, \mathfrak{p}) = \mathfrak{o}$, let s_{ij} be the least positive integer y such that

(6.3) $$\alpha_i^y \equiv \alpha_j^y \pmod{\mathfrak{p}}.$$

Then s_{ij} divides every other such y, and in particular the number $N(\mathfrak{p})-1 = p^t - 1$. Here $t \leqq k!$, the maximum possible degree of $\mathfrak{F}$.

Let m_p be the least common multiple of the numbers $p-1, p^2-1, \cdots, p^{k!}-1$ and if D has in all k distinct prime factors $p_1, p_2, \cdots, p_k$ let M be the least common multiple of $m_{p_1}, m_{p_2}, \cdots, m_{p_k}$. Then if *$s$ is chosen prime to both M and D* (and this choice can be made in an infinity of ways), $D^{(s)}/D$ *is prime to D*.

For if $(s, D)=1$ and $(D^{(s)}/D, D)\neq 1$, (6.1) holds. Then $s_{kl} \mid s$. Since $(s, M)=1$ and $s_{kl} \mid M$, $s_{kl}=1$ contradicting (6.11).

7. As in §6, let $p_1, \cdots, p_k$ be the distinct prime factors of D. By Theorems 4.5, 5.1 and Lemma 5.4, these primes are the only possible prime null divisors of (u) and its subsequences. Write

$$\Delta(u) = p_1^{r_1} \cdots p_k^{r_k} q, \qquad (q, D) = 1, \qquad r_i \geqq 0,$$

and let θ_i be the index of p_i in (u), where if p_i is not a null divisor, $\theta_i = 0$. By Lemma 5.3, $0 \leqq \theta_i \leqq r_i$, $(i = 1, 2, \cdots, k)$.

Now if R is the largest of $r_1, r_2, \cdots, r_k$, the numeric of p^{θ_i} is always less than k_{R^*}. Choose $s > k_R$ as in Theorem 6.1, and let (v) be the subsequence of (u) with general term $v_n = u_{ns}/u_s$ belonging to the polynomial $f^{(s)}(x)$. As in Theorem 6.1, let the discriminant of $f^{(s)}(x)$ be $D^{(s)}$. Then since $u_{ns} = \beta_1\alpha_1^{ns} + \cdots + \beta_k\alpha_k^{ns}$, we have by formula (5.3),

(7.1) $$\Delta(v) = \frac{\Delta(u)}{u_s^k}\,\frac{D^{(s)}}{D}.$$

Now $u_s \equiv 0 \pmod{p_i^{\theta_i}}$ and $(p_i, D^{(s)}/D) = 1$. Hence since $\Delta(v)$ is an integer, $\Delta(u) \equiv 0 \pmod{p_i^{k\theta_i}}$. Therefore $r_i \geqq k\theta_i$. If $\Delta(v) = p_1^{r_1'} \cdots p_k^{r_k'} q'$, $(q', D) = 1$, then $r_i' = r_i - k\theta_i$. Therefore

(7.2) $$r_i' < r_i \text{ if } \theta_i > 0; \; r_i' = r_i \text{ if } \theta_i = 0.$$

8. We now prove our second result indirectly. Suppose that the result is false. Then in any infinite set of normal divisibility sequences

$$\mathfrak{S}: \quad (u^{(1)}) = (u), (u^{(2)}), (u^{(3)}), \cdots, (u^{(m)}), (u^{(m+1)}), \cdots,$$

such that each sequence is a subsequence of its immediate predecessor, there must occur an infinity of non-primary sequences. Therefore there must exist a prime p dividing D which is a null divisor of an infinite number of the sequences $(u^{(m)})$. The general term of $(u^{(m+1)})$ is of the form $u_n^{(m+1)}=u_{ns_m}^{(m)}/u_{s_m}^{(m)}$, where the integer s_m specifies the particular subsequence of $(u^{(m)})$ selected. Consider now a set $\mathfrak{S}$ in which each $u^{(m)}$ satisfies the conditions imposed upon s in §6.

The considerations of the preceding section carry over to the relationship between $(u^{(m)})$ and $(u^{(m+1)})$. With an obvious extension of notation, let $\theta^{(m)}$ denote the index of p in $(u^{(m)})$ and p^{r_m} and $p^{r_{m+1}}$ the highest powers of p dividing $\Delta(u^{(m)})$ and $\Delta(u^{(m+1)})$. Then as in (7.2)

$$(8.1) \qquad r_{m+1} < r_m \text{ if } \theta^{(m)} > 0; \; r_{m+1} = r_m \text{ if } \theta^{(m)} = 0.$$

By our hypothesis, an infinite number of the $\theta^{(m)}$ are positive. But then (8.1) leads to an absurdity; for obviously $r=r_1\geqq r_2\geqq r_3\geqq \cdots \geqq 0$.

IV. Proof of third result

9. We assume as in the previous proofs that $D\neq 0$. In the Galois field $\mathfrak{F}$ of $f(x)$, a rational prime p which does not divide D remains unramified [9]. Accordingly the decomposition of p into prime ideal factors in $\mathfrak{F}$ is of the form

$$p = \mathfrak{p}_1\mathfrak{p}_2 \cdots \mathfrak{p}_l,$$

where the $\mathfrak{p}$ are all distinct.

Let σ_i be the least positive integer n such that

$$(9.1) \qquad \alpha_1^n \equiv \alpha_2^n \equiv \cdots \equiv \alpha_k^n \pmod{\mathfrak{p}_i} \qquad (i = 1, \cdots, l).$$

The restricted period τ of (u) modulo p is defined as the least value of n such that

$$u_{n+m} \equiv au_n \pmod{p} \qquad (m = 0, 1, 2, \cdots),$$

where a is some rational integer [5]. If p is prime to $\Delta(u)$, τ may be equally defined as the least positive integer n such that we have in $\mathfrak{F}$

$$\alpha_1^n \equiv \alpha_2^n \equiv \cdots \equiv \alpha_k^n \pmod{p}.$$

The following lemma therefore follows.

Lemma 9.1. *If p is a prime dividing neither $\Delta(u)$ nor D, then the restricted period τ of (u) modulo p is the least common multiple of the numbers $\sigma_1, \sigma_2, \cdots, \sigma_l$ associated with the congruence* (9.1) *above.*

10. Since $u_n=\beta_1\alpha_1 + \cdots +\beta_k\alpha_k^n$ and the α_i are distinct,

$$(10.1) \qquad \beta_i = u(\alpha_i)/f'(\alpha_i) \neq 0, \qquad (i = 1, \cdots, k).$$

Here $u(x)$ is the generator of the sequence (u) and $f'(x)$ the derivative of $f(x)$. Since $D=f'(\alpha_1)f'(\alpha_2)\cdots f'(\alpha_k)$, every prime ideal factor of the denominators of the β_i divides D. Let p be a rational prime. Since $\Delta(u)=\beta_1\beta_2\cdots\beta_k D=u(\alpha_1)u(\alpha_2)\cdots u(\alpha_k)$ we can state

LEMMA 10.1. *If $(p, D)=(p, \beta_1\beta_2\cdots\beta_k)=1$, then $(p, \Delta(u))=1$. Conversely if $(p, \Delta(u))=(p, D)=1$, then $(p, \beta_1\beta_2\cdots\beta_k)=1$, p a rational prime.*

Form the k sets of sums of the β taken $1, 2, \cdots, k$ at a time:

$$\beta_1+\beta_2+\cdots+\beta_i, \cdots, \beta_{k-i+1}+\beta_{k-i+2}+\cdots+\beta_k, \tag{10.2}$$

where $(i=1, 2, \cdots, k)$, and each set contains ${}_kC_i$ summands, not necessarily all distinct. Then take the symmetric products over each set

$$\mathrm{B}_i=\prod(\beta_1+\beta_2+\cdots+\beta_i) \qquad (i=1, 2, \cdots, k;\ \mathrm{B}_k=0).$$

Finally let $B=B(u)=\mathrm{B}_1\mathrm{B}_2\cdots\mathrm{B}_{k-1}$. Then B is a rational number of the form P/Q, where P and Q are integers, co-prime if $B\neq 0$ and the only primes dividing Q are divisors of D.

THEOREM 10.1. *If a prime p divides neither numerator nor denominator of B, its rank of apparition is the restricted period of (u) modulo p.*

By Lemma 10.1 any such prime p satisfies the hypothesis of Lemma 9.1. Let $\mathfrak{p}$ be any prime ideal factor of p, and ρ the rank of apparition of p. Since $p\mid u_\rho$ implies that $p\mid u_{n\rho}$,

$$\beta_1\alpha_1^{n\rho}+\beta_2\alpha_2^{n\rho}+\cdots+\beta_k\alpha_k^{n\rho}\equiv 0 \pmod{\mathfrak{p}} \qquad (n=0, 1, \cdots, k-1). \tag{10.3}$$

Since the β_i are integers modulo $\mathfrak{p}$ and prime to $\mathfrak{p}$, the determinant of this set of congruences is divisible by $\mathfrak{p}$. But this determinant is the difference product of the numbers $\alpha_1^\rho, \alpha_2^\rho, \cdots, \alpha_k^\rho$. Hence these numbers are not all distinct modulo p. *I say that*

$$\alpha_1^\rho\equiv\alpha_2^\rho\equiv\cdots\equiv\alpha_k^\rho \pmod{\mathfrak{p}}. \tag{10.4}$$

Otherwise the numbers can be grouped into two or more sets:

$$\alpha_{i_1}^\rho\equiv\alpha_{i_2}^\rho\equiv\cdots\equiv\alpha_{i_s}^\rho\equiv\zeta_1 \pmod{\mathfrak{p}}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_{j_1}^\rho\equiv\alpha_{j_2}^\rho\equiv\cdots\equiv\alpha_{j_t}^\rho\equiv\zeta_m \pmod{\mathfrak{p}}$$

such that the ζ are all distinct modulo p. The congruences (10.3) can then be replaced by

$$\beta_1'\zeta_1^n+\beta_2'\zeta_2^n+\cdots+\beta_m'\zeta_m^n\equiv 0 \pmod{\mathfrak{p}},$$

where the β' occur in the sets (10.2) of sums of β's. The determinant of the first m of these congruences as the difference product of the ζ is prime to $\mathfrak{p}$. Thus $\beta_1' \equiv \beta_2' \equiv \cdots \equiv \beta_m' \equiv 0 \pmod{\mathfrak{p}}$, so that $p \mid B$, contrary to hypothesis.

From (10.4) and the definition of the numbers σ in §9, we see that $\sigma \mid \rho$. Since this argument applies to all of the prime ideal factors of p, the least common multiple of $\sigma_1, \cdots, \sigma_l$ divides ρ. That is, by Lemma 9.1, $\tau \mid \rho$. But ρ always divides τ[2]. Hence $\rho = \tau$.

LEMMA 10.1. *If the number $B = B(u)$ is not zero, the rank of apparition of all save a finite number of primes in (u) is their restricted period.*

11. We now prove

THEOREM 11.1. *A sufficient condition that the number B be not zero is that the group of the characteristic polynomial of (u) be either alternating or symmetric.*

If B vanishes, one of the numbers of the set (10.2) vanishes. With a proper choice of notation we may assume that*

$$\beta_1 + \beta_2 + \cdots + \beta_i = 0, \qquad (k/2 \leqq i \leqq k). \tag{11.1}$$

We may also assume that $k > 4$, as the cases $k = 2, 3, 4$ may be easily discussed directly (see next theorem). Hence $i \geqq 3$.

If we represent the Galois group $\mathfrak{G}$ of $f(x)$ as a permutation group upon the k roots $\alpha_1, \cdots, \alpha_k$, then formula (10.1) shows that any permutation of the α induces the corresponding permutation upon the β. If $\mathfrak{G}$ is alternating or symmetric, it contains the permutation $S = (\alpha_1\alpha_{i+1})(\alpha_2\alpha_3)$. On applying S to (11.1), we obtain $\beta_{i+1} + \beta_2 + \beta_3 + \cdots + \beta_i = 0$. Hence $\beta_1 = \beta_{i+1}$. Similarly, $\beta_2 = \beta_{i+1}, \cdots, \beta_i = \beta_{i+1}$. Hence $\beta_{i+1} = 0$ contrary to (10.1).

The following result is proved by similar reasoning.

THEOREM 11.2. *For low orders of (u), sufficient conditions that $B(u) \neq 0$ are as follows:*

Order of (u)	*Condition of Galois group or characteristic polynomial*
2, 3	*none*
4	*order of group divisible by* 3
5	*$f(x)$ irreducible, or product of an irreducible quartic and linear factor*
6, 7	*group transitive and primitive.*

* It will be recalled that $\beta_1 + \beta_2 + \cdots + \beta_k = u_0 = 0$.

References

1. Annals of Mathematics, (2), vol. 18 (1916–17), pp. 51–64.
2. American Journal of Mathematics, vol. 58 (1936), pp. 577–584.
3. Duke Mathematical Journal, vol. 2 (1936), pp. 472–476.
4. These Transactions, vol. 35 (1933), pp. 600–628.
5. R. D. Carmichael, Quarterly Journal of Mathematics, vol. 48 (1920), pp. 343–372.
6. Bulletin de la Sciences Physiques Mathématiques de Kazan, (2), vol. 12 (1902), pp. 33–49. (In Russian.)
7. S. Lubelski, Crelle's Journal, vol. 102 (1930), pp. 66–67.
8. S. Lubelski, Prace Matematyczno-Fizyczne, vol. 43 (1936), p. 214.
9. Hilbert, *Die Theorie der algebraischen Zahlkörper*, Works, vol. 1, p. 85, Theorem 31; p. 144, Theorem 85.

California Institute of Technology,
Pasadena, Calif.

CONCERNING UNIQUENESS-BASES OF FINITE GROUPS WITH APPLICATIONS TO p-GROUPS OF CLASS 2*

BY

CHARLES HOPKINS†

A *uniqueness-basis* (U-basis) of a finite group G has been defined as "an ordered set of elements $Q_1, Q_2, \cdots, Q_\rho$ such that every element of G can be expressed uniquely in the form $Q_1^{x_1}Q_2^{x_2}\cdots Q_\rho^{x_\rho}$, where each x_i is a least positive residue modulo the order of Q_i."‡ In the case of the abelian groups the notion of U-basis is of undisputed importance: the theorem that every finite abelian group A has a U-basis may fairly be regarded as the cornerstone of the theory of abelian groups.

In the case of most non-abelian groups, however, the concept of U-basis is of doubtful advantage, especially in the general form above. Of greater usefulness, naturally, would be a "simplest type" of U-basis for the group under consideration. But the problem of constructing a definition of a "normal form" which shall be significant for reasonably general categories—the non-abelian p-groups, say—is an exceedingly difficult one. We offer a tentative definition in the case of the regular p-groups§ (§§2–3), which have, in common with the abelian p-groups, the property that the orders of the elements in every U-basis constitute a set of invariants of the group. In §4 we shall show how a "normal" U-basis may be used in constructing for every regular p-group G of class 2 a simply-isomorphic representation by l-matrices§ —matrices whose coordinates are residue classes modulo certain powers of p. These representations of G are of interest in that they usually involve a much smaller number of rows than do the matrix-representations whose coordinates are in a field. (The l-matrices are by no means novel; they have long been used for representing automorphisms of abelian p-groups.) In §5 we shall discuss the representation by l-matrices of the group of isomorphisms of G, and in §§6–7 we shall describe, very briefly, a representation of G as a multiplicative group in a finite ring.

1. In this section we shall state—for the most part without proof—several theorems which afford a set of criteria for the existence of a U-basis in a finite group G.

* Presented to the Society, September 5, 1936; received by the editors June 19, 1936.

† Corinna Borden Keen Research Fellow of Brown University.

‡ P. Hall, Proceedings of the London Mathematical Society, (2), vol. 36 (1934), p. 90.

§ Defined in §3.

THEOREM I. *For the ordered set of elements* $Q_1, Q_2, \cdots, Q_\rho$ *to constitute a U-basis for G it is necessary and sufficient that* (a) *each element of G be representable in the form* $Q_1^{x_1}Q_2^{x_2}\cdots Q_\rho^{x_\rho}$; (b) *the product of the orders of* $Q_1, Q_2, \cdots, Q_\rho$ *equal the order of G.**

Let $P_1, P_2, \cdots, P_n$ denote n operations of G whose orders are $g_1, g_2, \cdots, g_n$ respectively. Let P_x and P_y denote the products $P_1^{x_1}P_2^{x_2}\cdots P_n^{x_n}$ and $P_1^{y_1}P_2^{y_2}\cdots P_n^{y_n}$ respectively, $0 \leqq x_i < g_i$, $0 \leqq y_i < g_i$. We shall say that P_x and P_y are *formally distinct* if at least one x_k is not equal to y_k; we shall call them *effectively distinct* if they do not represent the same operation of G.

THEOREM II. *For the ordered set of elements* $P_1, P_2, \cdots, P_n$ *to constitute a U-basis for G it is necessary and sufficient that* (a) *the product* $g_1g_2\cdots g_n$ *equal the order of G*; (b) *any two formally distinct products* P_x *and* P_y *be effectively distinct.*

The following result is often useful:

THEOREM III. *If a finite group G, of order g, contains a set of subgroups* $G = G_1 \supset G_2 \supset \cdots \supset G_m$, *each* G_{i+1} *being a proper subgroup of index* g_i *in* G_i, *and if*

(a) G_m *contains a U-basis* $Q_1, Q_2, \cdots, Q_n$;

(b) $G_i - G_{i+1}$, $i = 1, 2, \cdots, m-1$, *contains an element* P_i *of order* g_i *such that* $P_i^{g_i}$ *is the lowest power of* P_i *which is in* G_{i+1}; *then the ordered set of elements* $P_1, P_2, \cdots, P_{m-1}, Q_1, \cdots, Q_n$ (*and* $Q_1, \cdots, Q_n, P_{m-1}, \cdots, P_2, P_1$ *as well*) *constitute a U-basis for G.*

By writing G_{m-1} in cosets with respect to G_m,

$$G_{m-1} = G_m + P_{m-1}G_m + \cdots + P_{m-1}^{g_{m-1}-1}G_m,$$

we see from Theorem I that $P_{m-1}, Q_1, \cdots, Q_n$ form a U-basis for G_{m-1}. The proof may be completed by induction.

THEOREM IV. *If a group G of order g contains two subgroups* H_1 *and* H_2, *of orders* h_1 *and* h_2 *respectively; and if*

(a) $h_1h_2 = g$;

(b) *the cross-cut* $H_1 \wedge H_2$ *is the identity*;

(c) H_1 *and* H_2 *contain the U-bases* $P_1, \cdots, P_\rho$ *and* $Q_1, \cdots, Q_\sigma$, *respectively*; *then the ordered set of elements* $P_1, \cdots, P_\rho, Q_1, \cdots, Q_\sigma$ (*or* $Q_1, \cdots, Q_\sigma, P_1, \cdots, P_\rho$) *constitute a U-basis for G.*

This theorem is easily proved by writing G in cosets with respect to H_1 (or H_2) and applying Theorem I.

* This rather obvious condition is mentioned by Hall for the case of a regular p-group; loc. cit., p. 95.

THEOREM V. *If G is the direct product of the subgroups $G_1, G_2, \cdots, G_m$, and if each subgroup G_i has the U-basis $Q_{i1}, \cdots, Q_{in_i}$, then a U-basis for G is given by the ordered set $Q_{11}, \cdots, Q_{1n_1}, Q_{21}, \cdots, Q_{2n_2}, \cdots$, etc.*

The wording of the theorem obviously implies that in this ordered arrangement the sets $(Q_{i1}, \cdots, Q_{in_i})$ may be permuted at will, provided that the sequence of the elements within the sets is undisturbed. For $m=2$ this theorem is a corollary of Theorem IV; by induction it can be proved for any m.

We conclude this section by giving several examples of groups which have uniqueness-bases.

A. *All dihedral groups.* For the dihedral group of order $2m$ (which is generated by two operations P and Q which satisfy the relations $P^2=Q^m=E$, $QP=PQ^{-1}$) the ordered set P, Q (and Q, P, as well) constitute a U-basis.

B. *Every symmetric group.* By Theorem III we may prove that the ordered set of cycles $a_1a_2, a_1a_2a_3, \cdots, a_1a_2 \cdots a_n$ constitute a U-basis for the symmetric group of degree n.

C. *Every alternating group $\mathcal{A}_n$.* The theorem is obvious for the alternating groups of degrees 2 and 3. We outline a proof by induction, assuming that the alternating group $\mathcal{A}_{n-1}$ of degree $n-1$ has a U-basis. There are three cases to consider: (a) when n is odd; (b) when n is divisible by 4; (c) when n is divisible by 2 and not by 4. In case (a) we know that $\mathcal{A}_n - \mathcal{A}_{n-1}$ contains the cycle $a_1a_2 \cdots a_n$. In case (b) it is easy to see that $\mathcal{A}_n$ contains the dihedral group of order n as a regular permutation group. In either case (a) or case (b), then, the proof may be completed by using Theorem IV. For case (c) we select from $\mathcal{A}_n$ the two permutations $s=(a_1a_n)(a_2a_3)$ and $t=(a_1a_2 \cdots a_{n/2})(a_{n/2+1} \cdots a_n)$. Now case (c) cannot arise for $n<6$, and it is easy to see that when $n \geqq 6$ all formally distinct products t^xs^y are effectively distinct, $0 \leqq x < n/2$; $0 \leqq y \leqq 1$; except for $x=0$ and $y=0$ the product t^xs^y will permute the letter a_n, and hence will not be a permutation in $\mathcal{A}_{n-1}$. One may now complete the proof by using Theorem II and the induction-hypothesis.

D. *The Sylow p-group $\sum_{p,n}$ of the general n-ary linear homogeneous group modulo p.* It is well known that $\sum_{p,n}$ can be represented by the group of matrices

$$(\alpha_{ij}) = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

where the $n(n-1)/2$ coordinates a_{ij} above the main diagonal are arbitrary elements of the Galois field of p elements. If we put $a_{in}=0, i=1, 2, \cdots, n-1$,

we obtain a representation of $\sum_{p,n-1}$; this group has only the identity in common with the abelian group defined by $a_{ij}=0$, $j\neq n$. By using Theorem IV and induction over n we may show that a U-basis for the group (α_{ij}) is given by the ordered set

$$E+e_{12},\ E+e_{13},\ E+e_{23},\ E+e_{14},\ \cdots,\ E+e_{1n},\ E+e_{2n},\ \cdots,\ E+e_{n-1,n},$$

where E is the n-rowed identity matrix and the e_{ij} are the usual basis-units of the n-ary matrix ring.

In conclusion, we offer the quaternion group as the simplest example of a group which has no U-basis.

2. In this section we introduce the notion of a ***normal*** U-basis. A U-basis $Q_1, Q_2, \cdots, Q_\rho$ of a finite group G is said to be ***normal*** with respect to G (in short, a *normal* U-basis) if for $i<j$; $i=1, 2, \cdots, \rho$; $j=2, 3, \cdots, \rho$, the elements Q_α satisfy the $\rho(\rho-1)/2$ equations

$$(1)\qquad Q_jQ_i = Q_iQ_j^{\beta_{jij}}Q_{j+1}^{\beta_{jij+1}}\cdots Q_\rho^{\beta_{ji\rho}}.\text{*}$$

As examples of groups having a normal U-basis we mention the dihedral groups and the groups $\sum_{p,n}$ above. Groups which contain a normal U-basis evidently constitute an exceptional category, as the restrictions imposed by the definition are relatively strong; for instance, each subgroup $\{Q_i, Q_{i+1}, \cdots, Q_\rho\}$ must be invariant in G, and G must be solvable.

An advantageous property of a normal U-basis is the following:

THEOREM I. *If a finite group G has a normal U-basis $Q_1, \cdots, Q_\rho$, the order of Q_i being g_i, then G is completely defined by the equations*

$$(2)\qquad Q_i^{g_i} = E$$

and the permutability relations (1) *above.*

Let $P_1, P_2, \cdots, P_\rho$ be a set of operations; i.e., elements which generate some group, and suppose that these elements are defined by equations (1) and (2) (assuming, of course, that we replace Q_i by P_i). Let $\overline{G}$ denote the group generated by $P_1, \cdots, P_\rho$. Since $\overline{G}$ and G are homomorphic under the correspondence defined by $P_i\sim Q_i$,† our theorem will follow if we can show that G and $\overline{G}$ have the same order; that is, if we can show that every product $\Pi=P_\alpha^{y_\alpha}P_\beta^{y_\beta}\cdots P_\alpha^{z_\alpha}\cdots$ of powers of $P_1, P_2, \cdots, P_\rho$ can be brought into the normal form $P_x=P_1^{x_1}P_2^{x_2}\cdots P_\rho^{x_\rho}$ by a finite number of reductions, each

* These equations define the commutators (Q_j, Q_i), $i<j$. From Theorem I below and the equation $(Q_i, Q_j)=(Q_j, Q_i)^{-1}$ it follows that (Q_i, Q_j) must have the form $Q_j^{\xi_{jij}}\cdots Q_\rho^{\xi_{ji\rho}}$.

† Burnside, *Theory of Groups of Finite Order*, 2d ed., p. 374.

reduction consisting of an interchange of two adjacent P's, using (1), followed by a reduction of exponents by means of (2). This is obviously possible when Π contains only $P_{\rho-1}$ and P_ρ as factors. The proof may be completed by induction: it is not difficult to show that a product Π involving no subscripts less than k can be brought into the normal form by a finite number of reductions, provided that this is true for all products containing subscripts greater than k.

Later we shall need the following generalization of the term "normal U-basis." Let G be a finite group and let Ψ be a group in which each element is an operator of G; suppose, further, that Ψ contains operators* which effect each of the inner isomorphisms of G. A U-basis $P_1, P_2, \cdots, P_\rho$ of G is said to be *normal with respect to* Ψ provided that

$$(3)\qquad P_i\psi = P_i^{\beta_{i\psi}} P_{i+1}^{\beta_{i+1\psi}} \cdots P_\rho^{\beta_{\rho\psi}}, \qquad i = 1, 2, \cdots, \rho,$$

where ψ is a variable operator in Ψ.

If for Ψ we take the group G itself, then this definition is equivalent to our earlier definition of a U-basis normal with respect to G. For in this case equations (3) contain the $\rho(\rho-1)/2$ equations

$$(4)\qquad P_i^{-1}P_jP_i = P_j^{\beta_{jij}} \cdots P_\rho^{\beta_{ji\rho}}, \qquad i < j.$$

THEOREM II. *If Ψ is of order p^t and if every element of G satisfies the equation $s^p = E$, then G contains a U-basis normal with respect to Ψ.*

By taking $\Psi \equiv G$ we have, as a corollary,

THEOREM III. *A finite p-group whose elements are of order p (identity excepted) contains a normal U-basis.*

In connection with Theorem III we observe that each exponent β_{jij} in (4) must be 1 modulo p; otherwise, the number of conjugates of P_j under P_i would contain a factor prime to p. Obviously Theorem III is not valid for p-groups in general.

Proof of Theorem II. We write s' for $s\psi$, where s and ψ are any elements of G and Ψ respectively. Let G_1 denote the subgroup of G which is generated by the totality of elements $c_1 = s^{-1}s'$. We define inductively the subgroup G_{i+1}. Suppose that G_i has already been defined, and suppose that c_i represents any element of G_i. Let c_i' and c_{i+1} denote $c_i\psi$ and $c_i^{-1}c_i'$ respectively, where ψ is any operator of Ψ. Then G_{i+1} is defined as the group generated by the totality of elements c_{i+1}.

* For a treatment of groups with operators see van der Waerden, *Moderne Algebra*, vol. I, p. 132. As concerns its effect on G, each operator ψ is equivalent to an operation T in the holomorph of G ($P_i\psi$ and $T^{-1}P_iT$ are the same element of G). It is hardly necessary to point out that Ψ need not be simply isomorphic with a subgroup of the holomorph of G.

Since each ψ effects a p-automorphism of G, we know that there is associated with a fixed operator ψ_λ a series of subgroups $G\supset G'\supset G''\supset\cdots\supset E$, each of index p in the preceding one, such that $s^{(\alpha)}\psi$, $s^{(\alpha)}$ being any element of $G^{(\alpha)}$, is equal to $s^{(\alpha)}$ multiplied by an element from $G^{(\alpha+1)}$.* From this we see at once that G_{i+1} is a proper subgroup of G_i, $i=1, 2, \cdots$; consequently, the series $G=G_0\supset G_1\supset G_2\supset\cdots$ must terminate in the identity E. Suppose that $G_f\equiv E$, but $G_{f-1}\not\equiv E$. We shall say that G is of class f with respect to Ψ. (By hypothesis, Ψ contains operators which bring about each of the inner isomorphisms of G. Hence G_1 contains the commutator subgroup of G. And if Ψ is G itself, then our definition of class coincides with the usual one.)

Now the group G_{f-1} of order $p^{n_{f-1}}$, say, is abelian and of type $1, 1, \cdots, 1$. For G_{f-1} we can construct a U-basis $P_{f-1,1}, P_{f-1,2}, \cdots$, and this U-basis will be normal with respect to Ψ, since every element of G_{f-1} satisfies the equation $c\psi=c$.†

If G is of class 1 with respect to Ψ, then our construction is at an end. Otherwise, we proceed by induction over G_i. Suppose that for G_{k+1} we have already constructed a U-basis normal with respect to Ψ. Now each quotient-group G_i/G_{i+1}, of order p^{n_i}, is abelian and of type $1, 1, \cdots, 1$. Hence we may construct for G_k/G_{k+1} a U-basis $u_1, u_2, \cdots, u_{n_k}$. From each of those cosets of G_k (with respect to G_{k+1}) which correspond to $u_1, u_2, \cdots$ we select an element as a representative, obtaining thereby n_k elements $P_{k\lambda}$, $\lambda=1, 2, \cdots, n_k$. The ordered set $P_{k1}, \cdots, P_{kn_k}$ followed by the elements of the U-basis for G_{k+1} (in the proper sequence) will constitute a U-basis for G_k which is normal with respect to Ψ. This assertion can readily be proved by using Theorem I of §1, together with the fact that the order of G_k is P^{r_k}, where r_k equals $\sum_{\alpha=k}^{f-1} n_\alpha$.

The construction which we have just given leads to a U-basis containing exactly m elements, where p^m is the order of G. It may be pointed out that every U-basis of G, normal or not, must contain exactly m elements. This follows from two considerations: every finite group whose elements are of order p, identity excepted, is a regular p-group;‡ the number of elements in a U-basis for a regular p-group is an invariant of the group.

3. In the introduction we mentioned an important category of p-groups, which resemble the abelian p-groups in that any two U-bases have the same number of elements of a given order. These are the regular p-groups, which have been defined in the following way:§ the p-group G will be called *regular*

* Miller, Blichfeldt, and Dickson, *Finite Groups*, p. 136.

† The proof of Theorem II depends mainly upon familiar properties of abelian groups A of order p^m and type $1, 1, \cdots, 1$: A has a U-basis, and the number of elements in every U-basis is exactly m.

‡ Hall, loc. cit., p. 74.

§ Hall, loc. cit., p. 73.

if, given any positive integer α and a pair of elements P and Q of G, it is always possible to find elements $S_3, S_4, \cdots, S_\rho$ all belonging to the commutator subgroup of $\{P, Q\}$ and satisfying the equation $(PQ)^{p^\alpha} = P^{p^\alpha} Q^{p^\alpha} S_3^{p^\alpha} S_4^{p^\alpha} \cdots S_\rho^{p^\alpha}$.

For an understanding of what follows, one must keep well in mind certain definitive properties of a regular p-group:

(a) The p^αth powers of the elements of G constitute a characteristic subgroup $\mho_\alpha(G)$.

(b) Those elements in G whose orders divide p^β constitute a characteristic subgroup $\Omega_\beta(G)$.

(c) The group G is conformal with an abelian group A.

(d) The orders of the elements in any U-basis of G are the same as the invariants $p^{\epsilon_1}, p^{\epsilon_2}, \cdots, p^{\epsilon_r}$ of A. (These orders have been called the *type-invariants* of G.)

(e) If Q_x is an element $Q_1^{x_1} Q_2^{x_2} \cdots Q_r^{x_r}$, written in the normal form with respect to the U-basis $Q_1, Q_2, \cdots, Q_r$, then the order of Q_x is equal to the order of its constituent $Q_\lambda^{x_\lambda}$, of highest order.

(f) If P and Q are any two elements of G, then the order of every element in the commutator subgroup of $\{P, Q\}$ divides the order of P (and of Q) relative to the central of G.

From (a) and (b) it is clear that G contains two series of characteristic subgroups: $G = \mho_0 \supset \mho_1 \supset \cdots \supset \mho_\delta = E$; $E = \Omega_0 \subset \Omega_1 \subset \cdots \subset \Omega_\delta = E$, where p^δ is the order of an element of highest order in G. From (c) it follows that δ is equal to the largest one of the ϵ's in (d).

(g) The order p^{ω_α} of $\Omega_\alpha/\Omega_{\alpha-1}$ equals the order of $\mho_{\alpha-1}/\mho_\alpha$, $\alpha = 1, 2, \cdots, \delta$. In particular, $G/\mho_1$ is of order ω_1. The ω's satisfy the inequalities $\omega_1 \geqq \omega_2 \geqq \cdots \geqq \omega_\delta$. Furthermore, ω_1 equals r, the number of type-invariants of G.

(h) For the ϵ's and ω's we have the relation $\sum_{i=1}^r \epsilon_i = \sum_{j=1}^r \omega_j = m$, where p^m is the order of G (and of A).

(i) $\mho_\alpha(G)$ and $\Omega_\beta(G)$ are conformal with $\mho_\alpha(A)$ and $\Omega_\beta(A)$ respectively, $\alpha, \beta = 0, 1, \cdots, \delta$.

Let $V_i(G)$ denote the cross-cut $\mho_1(G) \wedge \Omega_i(G)$, $i = 0, 1, \cdots, \delta$, and let $W_i(G)$ denote the group $\{V_i(G), \Omega_{i-1}(G)\}$.

(j) The groups $V_i(G)$, $W_i(G)$, $\Omega_i(G)/V_i(G)$, and $\Omega_i(G)/W_i(G)$ are conformal respectively with the groups which are obtained by replacing G with A.

(k) If the exponents of the invariants of A, arranged in descending order of magnitude, are given by $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$, and if in this arrangement the δ's constitute s sets, the jth set consisting of h_j equal δ's having the common

value e_i, then the order of $\Omega e_j(A)/We_j(A)$ is p^{h_j}. Obviously $m=\sum_{i=1}^{r}\delta_i=\sum_{j=1}^{s}h_je_j$.

Items (a) through (i) are given explicitly in the paper of Hall to which reference has already been made (see pp. 73–81); item (j) is contained implicitly in Hall's results; item (k) is a familiar result from the theory of abelian p-groups.

DEFINITION. *A U-basis $P_1, P_2, \cdots, P_r$ for a regular p-group G, the order of P_i being p^{δ_i}, is said to be ω-normal provided that*

(1) $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$;

(2) *when (P_j, P_i),* $i<j$, is expressed in the normal form*

$$(P_j, P_i) = P_1^{\alpha_{ji1}}P_2^{\alpha_{ji2}} \cdots P_r^{\alpha_{jir}}, \qquad i = 1, 2, \cdots, r; j = 2, 3, \cdots, r,$$

each α_{jik} is divisible by p for $k \leqq j$.

We state without proof two implications of this definition:

(3) for $k<j$, $\alpha_{jik}\equiv 0 \bmod p^{\delta_k-\delta_j}$ (see (e) and (f));

(4) if the normal form of (P_i, P_j), $i<j$, is given by $P_1^{\alpha_{ij1}} \cdots P_r^{\alpha_{ijr}}$, then the highest power of p that divides α_{ijk} also divides α_{jik}, and conversely. Moreover, (2) is clearly a consequence of (1) when no two of the δ's are equal.

For the abelian group A conformal with G any U-basis for which (1) is satisfied may be regarded as a "normal form," since all such U-bases are equivalent under the holomorph of A. In defining a normal form for a U-basis of G we must obviously demand more, and (2) seems to be the most natural additional requirement which can be satisfied in the case of every regular p-group. The qualifying phrase "ω-normal" is suggested by the fact that $P_1, P_2, \cdots, P_r$, regarded as representatives of a U-basis for $G/\mho_1$, constitute a *normal* U-basis for this quotient group; i.e.,

$$P_jP_i \equiv P_iP_jP_{j+1}^{\alpha_{jij+1}} \cdots P_r^{\alpha_{jir}} \bmod \mho_1.$$

As an attempt at defining a normal form for a U-basis, our definition above is obviously of no value unless we can prove the following theorem:

Every regular p-group G contains an ω-normal U-basis.

First, we explain a method for constructing a set of elements which satisfy requirements (1) and (2) above. To avoid repeated explanations, the symbols G, $\Omega(\ )$, m, h_j, e_j, etc., will have the same significance as in (a) through (k) above.

* This is the familiar notation for the commutator $P_j^{-1}P_i^{-1}P_jP_i$.

Let $\Omega_a(G)$ be the first term in the series $E=\Omega_0\subset\Omega_1\subset\cdots\subset\Omega_s=G$ for which W_a is a proper subgroup of Ω_a; i.e.,

(5) $$W_j\equiv\Omega_j,\quad j=0,1,\cdots,a-1;\qquad W_a\subset\Omega_a.$$

From (j) and (k) we see that the order of Ω_a/W_a is exactly p^{h_s}. Furthermore, every element of this quotient-group is of order p, except for the identity. Since Ω_a and W_a are characteristic subgroups of G, and since each element of G effects an automorphism of Ω_a/W_a, we can construct for this quotient-group a U-basis, $u_1, u_2, \cdots, u_{h_s}$, say, which is normal with respect to G (see Theorem II of §2). From the coset of Ω_a which corresponds to u_j we select any element Q_j as a representative, obtaining thereby the h_s elements $Q_1, Q_2, \cdots, Q_{h_s}$.

(6) Each $Q_j, j=1, 2, \cdots, h_s$, is of order p^a.

Otherwise, contrary to (5), we could find a $k<a$ for which $W_k\not\equiv\Omega_k$.

It is easy to see that every element of Ω_a can be expressed in the form $Q_1^{x_1}Q_2^{x_2}\cdots Q_{h_s}^{x_{h_s}}w_{a,x}$, where $w_{a,x}$ is an element in W_a. And since the Q_j's are representatives of a U-basis for Ω_a/W_a which is normal with respect to G, it is clear that for a variable element X in G we have the congruences

(7) $$X^{-1}Q_jX\equiv Q_jQ_{j+1}^{\xi_{j+1}}\cdots Q_{h_s}^{\xi_{h_s}}\ \mathrm{mod}\ W_a,\qquad j=1,2,\cdots,h_s.$$

From (i), (j), and (k) we also have the equality

(8) $$a=e_s.$$

If Ω_a is G itself, then our construction is at an end. If not, let Ω_b be the first term in the series $\Omega_{a+1}, \Omega_{a+2}, \cdots$ for which

(9) $$W_b\subset\Omega_b;\ W_{a+l}\equiv\Omega_{a+l},\qquad l=0,1,\cdots,b-a-1.$$

As above, we construct for Ω_b/W_b, which is necessarily of order $p^{h_{s-1}}$, a U-basis normal with respect to G, and from each coset of Ω_b which corresponds to one of these basis-elements we choose an element, obtaining the h_{s-1} elements $R_1, R_2, \cdots, R_{h_{s-1}}$.

(10) Each of the R's is of order p^b.

Suppose that one of them, R_λ, say, were of order less than p^b. Then R_λ would necessarily occur among the elements of a certain set Ω_k-W_k where $k<b$. From (5) and (9) we know that a is the only value of $k<b$ for which $W_k\subset\Omega_k$. And certainly R_λ cannot be an element in $\{\mho_1, \Omega_a\}$, since Ω_a is contained in W_b. As in (8) above, we have the equality $b=e_{s-1}$.

From the manner of their construction it follows that the R's satisfy congruences of the type

(11) $X^{-1}R_iX = R_iR_{i+1}^{\xi_{i+1}} \cdots R_{h_{s-1}}^{\xi_{h_{s-1}}} \bmod W_b, \qquad i = 1, 2, \cdots, h_{s-1},$

when X is any element of G.

Since W_b is a subgroup of $\{\mho_1, \Omega_a\}$, we may replace (11) by

(12) $X^{-1}R_iX \equiv R_iR_{i+1}^{\xi_{i+1}} \cdots R_{h_{s-1}}^{\xi_{h_{s-1}}}Q_1^{\eta_1} \cdots Q_{h_s}^{\eta_{h_s}} \bmod \mho_1.$

If Ω_b is not equal to G, then we continue the construction; and at this point it is reasonably clear how the construction advances. The final (the sth) stage will consist of selecting h_1 elements $P_1, P_2, \cdots, P_h$ of $\Omega_s \geqq G$ which correspond to a U-basis of G/W_s, this U-basis being, of course, normal with respect to G. Thus we obtain an ordered set of $r = \sum_{i=1}^{s} h_i$ elements

(13) $P_{11}, P_{12}, \cdots, P_{1h_1}, P_{21}, \cdots, P_{s-11}, \cdots, P_{s1}, \cdots, P_{sh_s},$

where $P_{s-1\,i}$ and P_{sj} denote the elements R_i and Q_j above.

For our purpose the significant properties of these elements are the two following: $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$, and

(14) $X^{-1}P_jX \equiv P_jP_{j+1}^{\xi_{j+1}} \cdots P_r^{\xi_r} \bmod \mho_1(G), \qquad j = 1, 2, \cdots, r.$

(For the sake of a simpler notation we have replaced P_{11} by P_1, P_{12} by $P_2, \cdots$, P_{sh_s} by P_r.) It is clear, therefore, that if the elements $P_1, P_2, \cdots, P_r$ form a U-basis for G, then this U-basis will be ω-normal.

To prove that $P_1, P_2, \cdots, P_r$ constitute a U-basis, it is sufficient to show that they form a canonical basis for G, since it is known that every canonical basis of a regular p-group is necessarily a U-basis.*

A canonical-basis of a regular p-group has been defined* as a set of $\omega(=\omega_1)$ elements $Q_1, Q_2, \cdots, Q_\omega$ (ω being the order of $G/\mho_1(G)$) which satisfy the following conditions:

(α) there exists a set of ω subgroups $G = K_1 \supset K_2 \supset \cdots \supset K_\omega \supset K_{\omega+1} = \mho_1$, each being invariant in G and a proper subgroup of the preceding, such that each of the ω sets $K_i - K_{i+1}$ contains exactly one of the Q_i's;

(β) the product of the orders of the Q_i's is as small as possible, consistent with (α).

It is known that

(γ) the product of the orders of the elements in and canonical basis must equal the order of G.†

To show that the elements $P_1, \cdots, P_r$ above form a canonical-basis, it is therefore sufficient to show that they satisfy requirements (α) and (γ).

* Hall, loc. cit., p. 91.

† Hall, loc. cit., p. 92.

Now (γ) is satisfied, since the elements (13) constitute s sets, the jth set containing h_j elements each of order p^{e_j} [see (k) and (8) above].

To prove that (α) is also satisfied, we observe from (14) that the group $F_i = \{P_i, P_{i+1}, \cdots, P_r, \mho_1\}$ is an invariant subgroup of G; moreover F_i is a proper subgroup of index p in F_{i-1}, and r is equal to ω. Hence the series $G = F_1 \supset F_2 \supset \cdots \supset F_r \supset \mho_1$ has the properties of the K-series in (α); and since $F_i - F_{i+1}$ contains P_i and no other one of the P's, it follows that requirement (α) is satisfied by the r elements $P_1, P_2, \cdots, P_r$.

Whenever an ω-normal U-basis is normal with respect to G, that is, whenever the congruences $(P_j,P_i) \equiv P_{i+1}^{\theta_{jii+1}} \cdots P_r^{\theta_{jir}} \bmod \mho_1$, $i<j$, can be replaced by equalities

$$(15) \qquad (P_j, P_i) = P_{i+1}^{\lambda_{jii+1}} \cdots P_r^{\lambda_{jir}},$$

then, from Theorem I of §2, we know that G is completely defined by the orders $p^{\delta_1}, \cdots, p^{\delta_r}$ and the exponents in (15). The existence of a G-normal U-basis is clearly exceptional and it is an open question whether the data provided by an ω-normal U-basis, namely, the orders of the elements and the $r(r-1)/2$ equations (2), are always sufficient to define the regular p-group from which they are derived.

A Note on l-Matrices.* Let (ξ_{ij}) be an r-rowed square matrix whose coordinates are arbitrary rational integers, and let $\delta_1, \delta_2, \cdots, \delta_r$ be a sequence of fixed positive integers satisfying the inequalities $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$. The l-matrix (x_{ij}) we shall define as the matrix formed by replacing each ξ_{ij} by the set of integers having the form $\xi_{ij} \pm \lambda_{ij} p^{\delta_j}$; that is, (x_{ij}) is the matrix whose jth column is composed of residue classes $[\xi_{ij}]$ modulo p^{δ_j}. The class $[\xi_{ij}]$ is characterized by the least positive residue of ξ_{ij} modulo p^{δ_j}; accordingly, we shall usually assume that the coordinate x_{ij} in the l-matrix (x_{ij}) is a least positive residue rather than a class $[\xi_{ij}]$. Two l-matrices are naturally to be regarded as distinct unless their corresponding coordinates are identical. The totality of distinct l-matrices constitute a set, which we shall designate by the expression $L_p(\delta_1, \delta_2, \cdots, \delta_r)$. The sum of two l-matrices (x_{ij}) and (y_{ij}) we shall define as the l-matrix (z_{ij}) for which z_{ij} is the least positive residue of $x_{ij}+y_{ij}$ modulo $p^{\delta_j}, j=1, 2, \cdots, r$; the product $(x_{ij})(y_{ij})$ is the l-matrix (w_{ij}) in which w_{ij} is the least positive residue of $\sum_{\alpha=1}^{r} x_{i\alpha} y_{\alpha j}$ modulo p^{δ_j}. Those l-matrices for which the conditions

(α) $x_{ij} \equiv 0 \bmod p^{\delta_j - \delta_i}, i>j,$

and

(β) $|x_{ij}| \not\equiv 0 \bmod p$

* These l-matrices were first defined, in a slightly different form, by A. Ranum, these Transactions, vol. 8 (1907), pp. 71–91.

hold constitute under multiplication a group, which we shall refer to as "the group of l-matrices."* This group, which we shall denote by the expression $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$, is simply isomorphic with the group of isomorphisms of the abelian p-group of type $\delta_1, \delta_2, \cdots, \delta_r$.*

Multiplication is not, in general, associative for any three matrices of the set $L_p(\delta_1, \cdots, \delta_r)$. For l-matrices (x_{ij}) which satisfy (α), however, multiplication is associative and distributive, and these l-matrices constitute a ring. Thus any expression $(a(x_{ij})+b(y_{ij})+ \cdots)^n(c(u_{ij})+d(v_{ij})+ \cdots)^m \cdots$, where a, b, m, n are positive integers, defines a unique l-matrix, and this consideration is the justification for our later notation. In particular, the l-matrix defined by the expression $a(x_{ij})$, where (x_{ij}) satisfies condition (α), may be regarded as $(x_{ij})+(x_{ij})+ \cdots +(x_{ij})$, where there are a terms, as $(\alpha_{ij})(x_{ij})$, or as $(x_{ij})(\alpha_{ij})$, where (α_{ij}) is the l-matrix whose diagonal elements are the least positive residues of a modulis $p^{\delta_1}, p^{\delta_2}, \cdots$, etc., the remaining elements being zeros.

4. This section is concerned with applying the data furnished by an ω-normal U-basis to the problem of constructing a one-to-one representation by l-matrices† for a special category of regular p-groups; that is, the regular p-groups of class 2.‡ The theory of representations of a group of order g by means of matrices with coefficients in a field of characteristic prime to g has been rather thoroughly exploited. Little is known, however, about representations of p-groups by matrices with coefficients in a field of characteristic p; and it is fair to say that the problem of representing a *given* p-group by l-matrices has received almost no attention.§

Since the group $GL_p(\delta_1, \cdots, \delta_r)$—the group of l-matrices—is simply isomorphic with the group of automorphisms of the abelian p-group of type $\delta_1, \delta_2, \cdots, \delta_r$, we can easily construct a representation of a p-group G_p if we can find an abelian p-group A_p which is transformed into itself by G_p. Thus we can always construct a multiply-isomorphic representation by l-matrices for any p-group G_p, whether regular or not, since G_p always contains invariant abelian subgroups. The real difficulty arises when we demand a 1–1 representation of G_p (that is, a representation which is simply isomorphic with G_p).

* Ranum, pp. 84–85.

† These l-matrices are defined in §3.

‡ Groups of class 2 (metabelian groups, in the terminology of American mathematicians) were originally defined as groups having abelian central quotient-groups (W. B. Fite, Proceedings of the American Association for the Advancement of Science, vol. 49 (1901), p. 41). They have also been defined as groups having abelian commutator subgroups. The two definitions are obviously equivalent.

§ The reciprocal problem, namely, the investigation of the subgroups of the group $GL_p(\delta_1,\delta_2, \cdots, \delta_r)$ has been widely discussed.

It is precisely these 1–1 representations which are of most interest, and in the case of regular p-groups of class 2 a method for constructing them may be developed from the theory of regular permutation groups.

Let G be a regular p-group of class 2 and of order p^m, $p>2$,* which is represented as a regular† permutation group on its p^m elements. Let $K(G)$ denote the holomorph of G, and let H denote that representation in $K(G)$ of the group of inner isomorphisms of G whose permutations omit the symbol for the identity of G. Let $s_1, s_2, \cdots, s_{p^m}$ denote the permutations of G, and let S_i denote that permutation of H which transforms G according to s_i. Since G is of class 2, its commutator subgroup $C(G)$ is contained in its central $\Gamma(G)$; furthermore, H is abelian, since it is simply isomorphic with G/Γ. From this we see that G is multiply isomorphic with H under the correspondence defined by $s_i \sim S_i^\lambda$, $i=1, 2, \cdots, p^m$, where λ is any fixed integer.

At this point we introduce several useful formulas, which one may readily verify:

(1) $(s_i, s_j) = (s_i^{-1}, s_j^{-1}) = (s_j^{-1}, s_i) = (s_j, s_i^{-1}) = (s_j, s_i)^{-1}$,

(2) $(s_i^x, s_j^y) = (s_i, s_j)^{xy}$,

(3) $(s_i, s_j) = (S_i, s_j) = (s_i, S_j)$; $(S_i, S_j) = E$.

Let p^μ be the order of the element of highest order in H. Since p is an odd prime $(p^\mu-1)/2$ is a positive integer; and we denote this integer by the letter a.

We shall need the following results:‡

(4) The p^m products $S_i^a s_i$ constitute a regular permutation group G_a which is abelian and conformal with G.

(5) The cross-cut $G \wedge G_a$ is the permutation group Γ.

(6) The group G_a is transformed into itself by G, and conversely.

Let $K(G_a)$ denote the holomorph of G_a, written as a permutation group on the letters of G, and let $I(G_a)$ be that representation in $K(G_a)$ of the group of isomorphisms of G_a which omits the symbol for the identity of G. Correspondingly, we define $I(G)$ as that representation in $K(G)$ of the group of isomorphisms of G which omits the symbol for the identity of G.

(7) The permutation group $I(G)$ is a subgroup of $I(G_a)$.

(8) Between the permutations $t_1, t_2, \cdots$ of G_a and those of G there is a

* The assumption $p>2$ is pertinent, since a group of order 2^m is regular only when it is abelian.

† The simultaneous occurrence of "regular" in two distinct and unrelated meanings is unfortunate; both usages, however, are already established in the literature. To avoid confusing repetitions of the adjective "regular", we agree that throughout the remainder of this section the symbol G shall be used precisely in the sense above.

‡ These Transactions, vol. 37 (1935), pp. 163–171. This paper will be referred to as H.

1–1 correspondence, which is defined by the equations $t_i = S_i^a s_i$ and $s_i = S_i^{-a} t_i$.

When, in the future, we speak of an element of G (or of G_a) as the "corresponding element" of an element in G_a (or in G), we shall have in mind the correspondence defined by the equations in (8).

Let $P_1, P_2, \cdots, P_r$, of orders $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$, be the permutations of an ω-normal U-basis for G; i.e., a set of permutations satisfying the conditions

$$(9) \qquad P_i^{p^{\delta_i}} = E, \qquad \delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r;$$

$$(10) \quad (P_j, P_i) = P_1^{\alpha_{ji1}} P_2^{\alpha_{ji2}} \cdots P_r^{\alpha_{jir}}, \quad i < j; i = 1, 2, \cdots, r; j = 2, 3, \cdots, r;$$

$$\alpha_{jik} \equiv 0 \bmod p, \text{ for } k \leqq j.$$

Let $\overline{S}_i$, $i = 1, 2, \cdots, r$ denote the permutation of H which transforms G according to P_i.

Of fundamental importance for our method of attack is the following result:*

(11) The permutations $A_1 = \overline{S}_1^a P_1$, $A_2 = \overline{S}_2^a P_2, \cdots, A_r = \overline{S}_r^a P_r$, which are necessarily of orders $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ respectively, constitute a U-basis for G_a.

In what follows we shall usually write T_i for $\overline{S}^{-a}$. From the equation

$$(12) \qquad P_i = T_i A_i$$

it follows that each basis-element P_i may be regarded as a product of a basis-element of G_a by an automorphism of G_a, and this is a point of view which we shall constantly emphasize.

From (1), (2), and (3) it is easy to establish the following formulas:

$$(13) \qquad (A_j, T_i) = (A_j, P_i) = (\overline{S}_j^a P_j, P_i) = (P_j, P_i)^{a+1};$$

$$(14) \qquad (A_i, T_j) = (A_j, T_i)^{-1}; (A_i, T_i) = E.$$

Now each T_i, $i = 1, 2, \cdots, r$, is a permutation in the group of isomorphisms of G_a (see (7) above). Regarded as an automorphism of G_a, each T_i is equivalent to a change of basis for G_a, defined by r equations of the form

$$(15) \qquad \begin{cases} A_1 \to A_1' = A_1^{1+\gamma_{1i1}} A_2^{\gamma_{1i2}} \cdots A_r^{\gamma_{1ir}} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ A_r \to A_r' = A_1^{\gamma_{ri1}} A_2^{\gamma_{ri2}} \cdots A_r^{1+\gamma_{rir}}, \end{cases}$$

where γ_{jik} is a least positive residue modulo p^{δ_k}.

If we denote the matrix of the exponents in (15) by $E_r + M_i$, E_r being the identity matrix of $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$, then the r matrices $E_r + M_i$ generate

* H, p. 177.

a 1–1 representation of H (and consequently a p^γ-1 representation of G, where p^γ is the order of the central $\Gamma(G)$). We know that each γ_{jik} in (15) is uniquely determined by the permutations T_i and $A_1, A_2, \cdots, A_r$. What is of equal interest, perhaps, is the fact that the γ_{jik} are uniquely determined by equations (9) and (10) together with the equations

$$(16)\qquad (P_j, P_i)P_k = P_k(P_j, P_i), \qquad i, j, k = 1, 2, \cdots, r,$$

for the reason that any r operations which satisfy (9), (10), and (16) generate a group which is simply isomorphic with G.*

We indicate a method for computing the γ_{jik} from the data in (9), (10), and (16). If $T_k^x = E$, $k=1, 2, \cdots, r$, where x ranges over all the exponents α_{jik} in (10), that is, if each constituent $P_k^{\alpha_{jik}}$ of (P_j, P_i) is in Γ, then it follows from (12) and (13) that γ_{jik} is equal to the least positive residue of $(a+1)\alpha_{jik}$ modulo p^{δ_k}. In general, however, it is impossible to find for G an ω-normal U-basis for which this favorable situation arises. The following procedure is always valid. In (13) we replace each P_k by T_kA_k, obtaining thereby a first approximation for (A_j, T_i) in terms of the basis elements of G_a:

$$(17)\qquad (A_j, T_i) = (P_1^{\alpha_{ji1}} \cdots P_r^{\alpha_{jir}})^{a+1}$$
$$= (T_1^{\alpha_{ji1}}A_1^{\alpha_{ji1}}T_2^{\alpha_{ji2}}A_2^{\alpha_{ji2}} \cdots T_r^{\alpha_{jir}}A_r^{\alpha_{jir}})^{a+1} = T_\alpha^{a+1}A_\alpha^{a+1}c_\alpha^{a+1},$$

where

$$T_\alpha = T_1^{\alpha_{ji1}} \cdots T_r^{\alpha_{jir}};\ A_\alpha = A_1^{\alpha_{ji1}} \cdots A_r^{\alpha_{jir}};\ c_\alpha = \prod_{k,l} (A_k, T_l)^{\alpha_{jik}\alpha_{jil}},$$
$$k < l;\ k = 1, 2, \cdots, r;\ l = 2, 3, \cdots, r.$$

Since (P_j, P_i) is in the central of G, we know that T_α must be the identity of H. We observe, in addition, that for $k \leqq j$ the order of $(A_k, T_l)^{\alpha_{jik}\alpha_{jil}}$ is less than the order of (A_k, T_l), since α_{jik} is divisible by p for $k \leqq j$; and for $k > j$, the first constituent of $(A_k, T_l) = P_1^{\alpha_{kl1}} \cdots P_r^{\alpha_{klr}}$ whose exponent is prime to p must have a subscript greater than l. Hence a finite number of reductions of the type (17) will suffice to bring (A_j, T_i) into the form $A_1^{\gamma_{ji1}} \cdots A_r^{\gamma_{jir}}$.

We have outlined a method for constructing, from the data of an ω-normal U-basis, a representation of G by a subgroup of the group $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$ of r-rowed l-matrices. Presently we shall extend this p^γ-1 representation of G to a 1–1 representation by imbedding each matrix $E_r + M_i$ in an $(r+1)$-

* The proof of this assertion is similar to the proof of Theorem I in §2. In interchanging the P's we make use of the formula $P_\beta^y P_\alpha^x = P_\alpha^x P_\beta^y (P_\beta, P_\alpha)^{xy}$, $\beta > \alpha$, which can be derived from (16); a basis for induction is provided by the fact that the exponent of each constituent P_λ in the normal form of (P_β, P_α) is divisible by p for $\lambda \leqq \beta$.

rowed l-matrix. First, however, we list several interesting properties of the matrices M_i:

(18) The highest power of p which divides γ_{jik} divides α_{jik}, and conversely.

(19) Every element in and below the main diagonal of each M_i is divisible by p.

(20) For $j>k$, γ_{jik} is divisible by $p^{\delta_k-\delta_j}$.

The truth of (18) follows from the details of (17) and from the fact that this reduction is reversible; i.e., we may reduce $(A_1^{\gamma_{ji1}}\cdots A_r^{\gamma_{jir}})^{-a-1}$ to the form $P_1^{\alpha_{ji1}}\cdots P_r^{\alpha_{jir}}$ if we replace A_k by $T_k^{-1}P_k$. Obviously (19) and (20) follow directly from (18), or we may derive (20) from the fact that the order of (A_k, T_i) must divide the order of A_k.

Since $(A_i, T_j)=(A_j, T_i)^{-1}$, we have the relation

$$\gamma_{ijk} = p^{\delta_k} - \gamma_{ijk}. \tag{21}$$

For $i=j$, this gives

$$\gamma_{jjk} = 0, \tag{22}$$

where this zero is the residue 0 modulo p^{δ_k}.

Let M_0 denote the r-rowed l-matrix in which each element in the jth column is the residue 0 modulo $p^{\delta_j}, j=1, 2, \cdots, r$. Let R_{li} denote the r-rowed l-matrix whose lth row is $\gamma_{li1}, \gamma_{li2}, \cdots, \gamma_{lir}$ and whose remaining rows contain only zeros. Now the least positive residues of the lth row in the product $R_{li}M_j$ are the exponents of the commutator $((A_l, T_i), T_j)$ (written, of course, in the normal form $A_1^{\lambda_1}A_2^{\lambda_2}\cdots A_r^{\lambda_r}$). Since this commutator is the identity of G_a, we see that $R_{li}M_j$ must equal M_0. But $M_i=R_{1i}+R_{2i}+\cdots+R_{ri}$.* This establishes an important property of the matrices M_i, namely,

$$M_iM_j \equiv M_0, \qquad i, j = 1, 2, \cdots, r.\dagger \tag{23}$$

For $i=j$, this gives

$$M_i^2 \equiv M_0. \tag{24}$$

Let H_M denote the group generated by the matrices $E_r+M_i, i=1, 2, \cdots, r$. As we have seen above, H_M is a 1–1 representation of H and a p^γ-1 representation of G. The element of H_M which corresponds to the "general" element $P_1^{x_1}P_2^{x_2}\cdots P_r^{x_r}$ of G is the l-matrix derived from $(E_r+M_1)^{x_1}(E_r+M_2)^{x_2}\cdots$

* See note on l-matrices in §3. The matrices M_i, R_{ji} are elements of a ring.

† We wish to emphasize the fact that M_iM_j is to be regarded not as the product of M_i and M_j in the ordinary sense, but as the l-matrix defined by this product.

$(E_r+M_r)^{x_r}$; and from (23) we see that this product can be represented in the simple form*

$$(25) \qquad (E_r + M_1)^{x_1} \cdots (E_r + M_r)^{x_r} \equiv E_r + \sum_{k=1}^{r} x_k M_k.$$

We shall now construct a 1–1 representation of G as a subgroup of the group $GL_p(\delta_0, \delta_1, \delta_2, \cdots, \delta_r)$, where δ_0 is any fixed integer not less than δ_1. First, we define M_i' as the $(r+1)$-rowed l-matrix

$$\begin{pmatrix} 0 & 0 & 0 \cdots & 0 \\ 0 & \gamma_{1i1} & \gamma_{1i2} \cdots & \gamma_{1ir} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \gamma_{ri1} & \gamma_{ri2} \cdots & \gamma_{rir} \end{pmatrix}, \qquad i = 1, 2, \cdots, r,$$

where the elements in the first column are the residues 0 modulo p^{δ_0}. In order to avoid altering much of our earlier notation, we shall number the rows (and columns) in M_i' (and in the other $(r+1)$-rowed matrices which we shall presently define) by the sequence $0, 1, 2, \cdots, r$. Let E' be the identity matrix† of $GL_p(\delta_0, \delta_1, \cdots, \delta_r)$. It is at once evident that the matrices $E'+M_i'$ generate a group G_M which is simply isomorphic with H_M.

We denote by L_i' the $(r+1)$-rowed l-matrix which has in row 0 and column i the residue 1 modulo p^{δ_i} and zeros elsewhere. We shall denote the sum $M_i'+L_i'$ by the symbol N_i'.

The main result of this section is the following:

THEOREM I. *The r matrices $E'+N_i'$ generate a group G_N which is simply isomorphic with G under the correspondence defined by $P_i \sim E'+N_i'$, $i=1, 2, \cdots, r$.*

In proving this theorem we shall make use of the following known result: If $B_1, B_2, \cdots, B_r$, of orders $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ respectively, constitute a U-basis for an abelian group of order p^m, and if $\Theta_1, \Theta_2, \cdots, \Theta_r$ are a set of automorphisms of this group, each being defined by the r equations

$$(\alpha) \qquad (B_j, \Theta_i) = B_1^{\gamma_{ji1}} \cdots B_r^{\gamma_{jir}},$$

where the δ's and the γ_{jik} are the same as in (9) and (15) above, then the p^m products $\Theta_i B_i$, $i=1, 2, \cdots, r$, generate a group which is simply isomorphic with G under the correspondence $P_i \sim \Theta_i B_i$.‡

* That is, both sides of this equation define the same l-matrix of H_M.

† At this point we shall drop the subscripts from the identity matrices.

‡ This result is contained implicitly in the paper on metabelian groups which has been quoted above (see H, p. 193). It is proved there that the p^m products $\Theta_1^{x_1} \cdots \Theta_r^{x_r} B_1^{x_1} \cdots B_r^{x_r}$, $0 \leqq x_i < p^{\delta_i}$, constitute a group simply isomorphic with G; and it is easy to see that one may bring into this form any product $(\Theta_\alpha B_\alpha)^{\nu_\alpha} \cdots (\Theta_\beta B_\beta)^{\nu_\beta} \cdots$.

To prove Theorem I it is therefore sufficient to show that

(i) $E'+N_i'$ equals U_iV_i, where U_i and V_i denote $E'+M_i'$ and $E'+L_i'$, respectively;

(ii) $V_1, V_2, \cdots, V_r$ are of orders $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ respectively, generate an abelian group, and constitute a U-basis for this group;

(iii) the U_i and V_j satisfy the equations

$$(\beta) \qquad (V_j, U_i) = V_1^{\gamma_{ji1}} \cdots V_r^{\gamma_{jir}}, \qquad i < j;\ i = 1, \cdots, r;\ j = 2, \cdots, r.$$

First, we write several useful formulas, which can be verified from the rules for multiplying l-matrices:*

$$(26) \qquad \begin{aligned} &L_iL_j \equiv M_0; M_iM_j \equiv M_0 \quad ((23) \text{ above}); \quad M_iL_j \equiv M_0;\\ &L_jM_i \equiv \gamma_{ji1}L_1 + \cdots + \gamma_{jir}L_r, \qquad i, j = 1, 2, \cdots, r. \end{aligned}$$

$$(27) \qquad \begin{aligned} &U_i^{x_i} = (E + M_i)^{x_i} \equiv E + x_iM_i; \quad V_j^{y_j} = (E + L_j)^{y_j} \equiv E + y_jL_j;\\ &U_1^{x_1} \cdots U_r^{x_r} \equiv E + \sum_{k=1}^{r} x_kM_k; \qquad V_1^{y_1} \cdots V_r^{y_r} \equiv E + \sum_{k=1}^{r} y_kL_k. \end{aligned}$$

To prove (i) we have

$$U_iV_i = (E + M_i)(E + L_i) \equiv E + M_i + L_i + M_iL_i \equiv E + M_i + L_i = E + N_i.$$

(By definition, $N_i = L_i + M_i$; from (26), $M_iL_i \equiv M_0$.)

We now prove (ii). The equation $V_i^{y_i} = E$ requires $E + y_iL_i \equiv E$. Since p^{δ_i} is the smallest value of y_i for which $y_iL_i \equiv M_0$, it follows that the order of V_i is exactly p^{δ_i}. The permutability of V_i and V_j follows from (26). For $V_1, V_2, \cdots, V_r$ to constitute a U-basis for the abelian group which they generate, it is sufficient that the equation $V_1^{y_1} \cdots V_r^{y_r} = E$ be satisfied only by $y_i \equiv 0 \bmod p^{\delta_i}$. That this is the case follows directly from (27) and the linear independence of the L's.

Finally we prove (iii). From (26) and (27) we derive the equalities†

$$\begin{aligned} (V_j, U_i) = V_j^{-1}U_i^{-1}V_jU_i &\equiv (E - L_j)(E - M_i)(E + L_j)(E + M_i)\\ &\equiv (E - L_j - M_i + L_jM_i)(E + L_j + M_i + L_jM_i)\\ &\equiv E + \sum_{k=1}^{r} \gamma_{jik}L_k\\ &\equiv \prod_{k=1}^{r} (E + L_k)^{\gamma_{jik}} = V_1^{\gamma_{ji1}}V_2^{\gamma_{ji2}} \cdots V_r^{\gamma_{jir}}. \end{aligned}$$

* We shall drop all primes, since from this point on we shall deal exclusively with $(r+1)$-rowed matrices. Note that M_0 is now the $(r+1)$-rowed null-matrix of the set $L_p(\delta_0, \delta_1, \cdots, \delta_r)$.

† It is understood, of course, that the notation $E - L_j$ is merely a convenient substitute for $E + (p^{\delta_j} - 1)L_j$.

This completes the determination of our 1–1 representation of G by means of l-matrices. As we have seen, this particular representation G_N is completely defined by the orders of the basis elements $P_1, P_2, \cdots, P_r$ of G, the exponents α_{jik} in the $r(r-1)/2$ equations (10), and the permutability relation $(P_i, P_j)P_k = P_k(P_i, P_j)$. In respect to the totality of possible representations of G by l-matrices $L_p(\delta_1, \delta_2, \cdots, \delta_r)$, the representation G_N may be regarded as a normal form, in that for every matrix in G_N the elements in and below the main diagonal are congruent modulo p to 1 and 0 respectively.

We observe that the matrices $E+\sum_{k=1}^r x_k L_k$, $0 \leqq x_k < p^{\delta_k}$, define a 1–1 representation of the abelian group G_a, and the corresponding elements of G_N are the l-matrices derived from the products $(E+\sum_{k=1}^r x_k M_k)(E+\sum_{k=1}^r x_k L_k) \equiv E+\sum_{k=1}^r x_k N_k$ (see (8) and (25) above). Thus we obtain all the elements of G_N from a "general" matrix, whose coordinates are linear functions of r parameters, by specializing these parameters and taking least positive residues. We shall see, later, that the N's combine under multiplication according to the linear formula $N_i N_j \equiv \sum_{k=1}^r \gamma_{ijk} N_k$.

5. The notation which we shall use in this section is that of the preceding. Our objective is to characterize those matrices of the group $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$ which represent automorphisms* of G. We know, of course, that the group of isomorphisms of G_a is simply isomorphic with $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$; and from (7) of §4 it follows that $I(G)$ is simply isomorphic with a certain subgroup of $GL_p(\delta_1, \cdots, \delta_r)$. Now every matrix $X=(x_{ij})$ in $GL_p(\delta_1, \cdots, \delta_r)$ is characterized by two conditions on the coordinates x_{ij}, namely,

$$x_{ij} \equiv 0 \bmod p^{\delta_j-\delta_i} \text{ for } i > j; \tag{1}$$

$$| x_{ij} | \not\equiv 0 \bmod p. \tag{2}$$

Our problem, therefore, is to determine the additional restrictions which must be imposed on the coordinates x_{ij} in order that x shall define an automorphism of G. It is possible to determine these additional conditions as congruences involving the α_{jik} by regarding X as the coefficient-matrix in the correspondence

$$\begin{cases} P_1 \sim P_1^{x_{11}} \cdots P_r^{x_{1r}} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ P_r \sim P_1^{x_{r1}} \cdots P_r^{x_{rr}}. \end{cases}$$

It is much easier, however, to determine them in terms of the γ_{jik}, and this is

* Throughout this article the term "automorphism" of G denotes a 1-1 isomorphism of G with itself.

the procedure which we shall adopt. First, we regard X as the coefficient-matrix of the "general" automorphism of G_a, defined by the correspondence

$$\left\{\begin{array}{l} A_1 \sim B_1 = A_1^{x_{11}} \cdots A_r^{x_{1r}} \\ \cdots\cdots\cdots\cdots\cdots \\ A_r \sim B_r = A_1^{x_{r1}} \cdots A_r^{x_{rr}}, \end{array}\right. \tag{3}$$

the x_{ij} being least positive residues.

Let π_x denote that permutation of $I(G_a)$ which is defined by X. Now π_x transforms G either into itself or a conjugate permutation group in $K(G_a)$; furthermore, G is generated by the r permutations $P_i = T_i$, $i = 1, \cdots, r$ (see (11) and (12) of §4). Hence, for X to define an automorphism of G it is necessary and sufficient that $\pi_x^{-1}T_iA_i\pi_x$, $x = 1, 2, \cdots, r$, be a permutation of G. Now the permutation of G which corresponds to the permutation $A_z = A_1^{z_1} \cdots A_r^{z_r}$ of G_a is the product T_zA_z, where T_z is $T_1^{z_1} \cdots T_r^{z_r}$ (see (8) of §4). Consequently, for X to define an automorphism of G it is necessary and sufficient that

$$\pi_x^{-1} T_i \pi_x \text{ be equal to } T_1^{x_{i1}} \cdots T_r^{x_{ir}}, \qquad i = 1, 2, \cdots, r. \tag{4}$$

The sufficiency of this condition is obvious. That it is necessary, as well, follows from the fact that if we multiply a given permutation $A_z = A_1^{z_1} \cdots A_r^{z_r}$, $0 \leqq z_i < p^{\delta_i}$, on the left by all the permutations of H, then only one of these products will be a permutation in G; that is, the multiplying permutation T_z is unique, although its representation in the form $T_1^{z_1} \cdots T_r^{z_r}$, $0 \leqq z_i < p^{\delta_i}$, is never unique.

We seek a more explicit form for condition (5). As in the preceding section, each T_i is equivalent to a substitution

$$\left\{\begin{array}{l} A_1 \to A_1' = A_1^{1+\gamma_{1i1}} \cdots A_r^{\gamma_{1ir}} \\ \cdots\cdots\cdots\cdots\cdots \\ A_r \to A_r' = A_1^{\gamma_{ri1}} \cdots A_r^{\gamma_{rir}}, \end{array}\right. \tag{5}$$

whose matrix we have already denoted by $E + M_i$. Similarly, the product $T_1^{x_{i1}} \cdots T_r^{x_{ir}}$ is equivalent to a substitution on the A's whose matrix is given by $E + \sum_{k=1}^{r} x_{ik}M_k$ (see (25) of §4). And the matrix of that substitution (on the A's) which is equivalent to the permutation $\pi_x^{-1}T_i\pi_x$ is clearly equal to $X^{-1}(E + M_i)X$. Hence (5) is equivalent to the equations

$$E + \sum_{k=1}^{r} x_{ik}M_k \equiv X^{-1}(E + M_i)X, \qquad i = 1, \cdots, r; \tag{6}$$

and these, in turn, are equivalent to the set

$$\sum_{k=1}^{r} x_{ik}M_k \equiv X^{-1}M_iX. \tag{7}$$

The main result of this section may be expressed by the theorem:

The group of isomorphisms of G is simply isomorphic with the group generated by those l-matrices $L_p(\delta_1, \cdots, \delta_r)$ for which the following conditions are satisfied:

(a) $$x_{ij} \equiv 0 \bmod p^{\delta_j - \delta_i} \quad \textit{for} \quad i > j;$$

(b) $$| x_{ij} | \not\equiv 0 \bmod p;$$

(c) $$\sum_{k=1}^{r} x_{ik}M_k \equiv X^{-1}M_iX, \qquad i = 1, 2, \cdots, r.$$

In conclusion, we state a useful relation, namely,

$$\sum_{k=1}^{r} \gamma_{ijk}M_k \equiv M_0, \qquad i = 1, 2, \cdots, r, \tag{8}$$

which can be derived by substituting $E+M_i$ for X in (7) above.

6. In this section we shall investigate the abstract structure of the representation G_N, and we shall see that the matrices $E, N_1, \cdots, N_r$ may be regarded as basis-units in a certain finite ring.

We have already pointed out that the general element of G_N is obtained by reducing the matrix $J_x = E + \sum_{k=1}^{r} x_k N_k$, where each x_k ranges from 0 to $p^{\delta_k} - 1$. Since the product $J_z = J_x J_y$ must occur in the form $J_z = E + \sum z_k N_k$, each of the r^2 products $N_i N_j$ must obviously be equivalent to a linear function of the matrices $E, N_1, \cdots, N_r$. This linear relation is given by the formula

$$N_iN_j \equiv \sum_{k=1}^{r} \gamma_{ijk}N_k. \tag{1}$$

In deriving this formula we replace N_i by $L_i + M_i$.* From (26) of §4 and (8) of §5 we have the chain of equations

$$N_iN_j = (L_i + M_i)(L_j + M_j) \equiv L_iM_j \equiv \sum_{k=1}^{r} \gamma_{ijk}L_k \equiv \sum_k \gamma_{ijk}L_k + \sum_k \gamma_{ijk}M_k$$
$$\equiv \sum_k \gamma_{ijk}(L_k + M_k) \equiv \sum_{k=1}^{r} \gamma_{ijk}N_k.$$

* Observe that M_i is the $(r+1)$-rowed matrix M_i' of §4. It is evident that (8) above is valid if we replace M_i by M_i'.

Since $\gamma_{jik} \equiv -\gamma_{ijk}$ mod p^{δ_k} and $\gamma_{iik} \equiv 0$ mod p^{δ_k} (see (21) and (22) of §4), we have for the N's the further relations

$$N_iN_j \equiv -N_jN_i \tag{2}$$

(the interpretation of this congruence is obvious);

$$N_i^2 \equiv N_0, \tag{3}$$

where N_0 is the $(r+1)$-rowed null matrix.

From (1) and (2) it is easy to show that the product of any three of the N's is the null matrix; that is,

$$N_iN_jN_k \equiv N_0, \qquad i, j, k = 1, 2, \cdots, r. \tag{4}$$

At this point it is fairly evident that with the group G_N there is associated a finite ring having as basis-units the matrices $E, N_1, \cdots, N_r$. We wish to show that this ring can be constructed without assuming the existence of G_N.

We start with a system $\mathfrak{S}$ of double composition in which all the ring postulates are satisfied except (possibly) associativity of multiplication. We designate a set of $r+1$ linearly independent basis-units for $\mathfrak{S}$ by $v_e, v_1, \cdots, v_r$, and we assume that every element of $\mathfrak{S}$ can be represented uniquely in the form $v_x = x_0v_e + \sum_{i=1}^{r} x_iv_i$, where the x_i are arbitrary rational integers.

We assume that multiplication for the basis-units (and accordingly for every element of $\mathfrak{S}$) is defined by the equations

$$\left\{\begin{aligned} v_e^2 &= v_e, \\ v_ev_i &= v_iv_e = v_i, \qquad\qquad i = 1, 2, \cdots, r, \\ v_iv_j &= \sum_{k=1}^{r} \gamma_{ijk}v_k,^* \end{aligned}\right. \tag{5}$$

where the γ_{ijk} have the same values as in (1) above. We recall that each γ_{ijk} is a positive integer less than p^{δ_k}, and that these r^3 integers satisfy

$$\gamma_{ijk} \equiv -\gamma_{jik} \bmod p^{\delta_k} \qquad [\S4, (21)] \tag{6}$$

$$\left.\begin{aligned} \gamma_{ijk} &\equiv 0 \quad \bmod p^{\delta_k-\delta_i} \text{ for } i > k \\ \gamma_{ijk} &\equiv 0 \quad \bmod p^{\delta_k-\delta_j} \text{ for } j > k \end{aligned}\right\} \ [\S4, (20) \text{ and } (21)] \tag{7}$$

$$\sum_{\alpha=1}^{r} \gamma_{ij\alpha}\gamma_{\alpha kl} \equiv 0 \qquad \bmod p^{\delta_l} \qquad [\S4, (23)] \tag{8}$$

$$\sum_{\alpha=1}^{r} \gamma_{jk\alpha}\gamma_{i\alpha l} \equiv 0 \qquad \bmod p^{\delta_l} \qquad [\S5, (8)]. \tag{9}$$

* It is known, of course, that the various assumptions above are always consistent. In fact, a system of the type $\mathfrak{S}$ exists if we replace the γ_{ijk} in (5) by r^3 arbitrary integers; and if multiplication is associative (which is generally not the case), then this system is a ring of rank $r+1$, having the ring of integers as its coefficient-domain.

Now those elements v_x for which x_k is divisible by p^{δ_k}, $k=0, 1, \cdots, r$,* constitute under addition a modulus $\mathfrak{M}$. We wish to show that the elements of this modulus constitute an invariant ideal in $\mathfrak{S}$. Since they form a modulus, it is sufficient to show that $v_x\alpha_y$ and $\alpha_y v_x$ are each of the form $\lambda_0 p^{\delta_0}v_e+\sum_{i=1}^{r}\lambda_i p^{\delta_i}v_i$, where v_x and α_y are any elements of $\mathfrak{S}$ and $\mathfrak{M}$ respectively.

From (5) we have†

$$\left(\sum_{i=1}^{r} x_i v_i\right)\left(\sum_{j=1}^{r} y_j p^{\delta_j} v_j\right) = \sum_{k=1}^{r}\left[\sum_{i=1}^{r} x_i\left(\sum_{j=1}^{r} y_j p^{\delta_j}\gamma_{ijk}\right)\right] v_k.$$

Since $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$, it is obvious that when k is at least equal to both i and j, then the coefficient of v_k is divisible by p^{δ_k}. From (7), however, we know that γ_{ijk} has the form $\gamma'_{ijk}p^{\delta_k-\delta_i}$ (or $\gamma''_{ijk}p^{\delta_k-\delta_j}$) for $k<i$ (for $k<j$). In every case then, the coefficient of v_k is divisible by p^{δ_k}, $k=0, 1, \cdots, r$. Similarly, we may show that $\alpha_y v_x$ is an element of $\mathfrak{M}$. We denote this invariant ideal by the letter $\mathfrak{I}$.

From a familiar result in the theory of ideals we know that the residue classes of $\mathfrak{S}$ with respect to $\mathfrak{I}$ form a system $\overline{\mathfrak{S}}$ which is homomorphic with $\mathfrak{S}$. If we denote by $u_e, u_1, \cdots, u_r$ the elements of $\overline{\mathfrak{S}}$ to which $v_e, v_1, \cdots, v_r$ correspond respectively in this homomorphism, then it is easy to see that the u's constitute a basis for $\overline{\mathfrak{S}}$, and that every element of $\overline{\mathfrak{S}}$ can be represented in the form $u_x=[x_0]u_e+\sum_{i=1}^{r}[x_i]u_i$, where $[x_i]$ is the symbol for the class of integers congruent to x_i modulo p^{δ_i}. In view of the homomorphism above, we know that the sum u_x+u_y is represented by $[x_0+y_0]u_0+\sum_i[x_i+y_i]u_i$, while the multiplication table for the u's is given by

$$\begin{aligned} u_e^2 &= u_e, \\ u_e u_i &= u_i u_e = u_i, \qquad i = 1, 2, \cdots, r, \\ u_i u_j &= \sum_{k=1}^{r} \gamma_{ijk} u_k. \end{aligned} \tag{10}$$

The condition for associativity of multiplication in $\overline{\mathfrak{S}}$ is given by

$$\left[\sum_{\alpha=1}^{r}[\gamma_{ij\alpha}]_\alpha[\gamma_{\alpha kl}]_l\right]_l = \left[\sum_{\alpha=1}^{r}[\gamma_{jk\alpha}]_\alpha[\gamma_{i\alpha l}]_l\right]_l.‡$$

But from (8) and (9) we know that both sides of this equation are equal to the residue class 0 modulo p^{δ_l}. Hence $\overline{\mathfrak{S}}$ is a ring, since it is homomorphic

* These p^{δ_k} are the type-invariants of G [see §4].

† Obviously $v_x \cdot y_0 p^{\delta_0} v_e$ is in $\mathfrak{M}$.

‡ The symbol $[\xi]_\lambda$ denotes the class of integers having the form $\xi \pm np^{\delta_\lambda}$. We need consider only elements in $\overline{\mathfrak{S}}$ of the form $\sum_{i=1}^{r}[x_i]u_i$, since $(u_e u_x)u_y$ is obviously equal to $u_e(u_x u_y)$.

with a system $\mathfrak{S}$ which satisfies every ring postulate except associativity of multiplication.

From (6), which implies that $[\gamma_{ijk}]_k = [-\gamma_{jik}]_k$, we may easily show that those elements in $\bar{\mathfrak{S}}$ which have the form $u_x' = \sum_{i=1}^r [x_i]u_i$ constitute a nilpotent* ring $\mathfrak{R}$ of index 3 in which the square of every element is the zero element. It is clear that $\mathfrak{R}$ is an invariant subring of $\bar{\mathfrak{S}}$; furthermore, the elements of $\mathfrak{R}$ constitute under addition an abelian group A_R which is simply isomorphic with G_a of §4. One may also show that the elements of $\bar{\mathfrak{S}}$ having the form $u_x'' = [1]u_e + u_x'$ constitute under multiplication a group G_R of order p^m and class 2. That they form a group is obvious, since $u_x''u_y''$ can be brought into the form $[1]u_e + u_z'$, while every u_x'' has an inverse, namely, $[1]u_e + \sum_{i=1}^r [-x_i]u_i$. The commutator (u_i'', u_j''), where $u_i'' = u_e + u_i$, is $u_e + \sum_{\alpha=1}^r 2\gamma_{ij\alpha}u_\alpha$, and by using (8) and (9) one readily computes for $(u_i'', u_j'')u_k''$ and $u_k''(u_i'', u_j'')$ the common value $u_e + u_k + \sum_{\alpha=1}^r 2\gamma_{ij\alpha}u_\alpha$. Since G_R is generated by the u_k'''s, its commutator subgroup must be contained in its central; hence G_R is of class 2.

Since $\bar{\mathfrak{S}}$ contains a principal unit u_e, we can construct a 1–1 matrix-representation of $\bar{\mathfrak{S}}$ by the familiar process of post-multiplication. We denote the products $u_e(u_e+u_i)$, $u_1(u_e+u_i)$, $\cdots$, $u_r(u_e+u_i)$ by $\bar{u}_e$, $\bar{u}_1$, $\cdots$, $\bar{u}_r$ respectively. Then the matrix-representation of the element u_e+u_i is the coefficient-matrix of the linear transformation

$$\left\{\begin{array}{lll} \bar{u}_e = u_e & & + \, u_i \\ \bar{u}_1 = & (1+\gamma_{1i1})u_1 + \cdots + \gamma_{1ir}u_r & \\ \cdots & \cdots\cdots\cdots\cdots\cdots & \\ \bar{u}_r = & \gamma_{ri1}u_1 + \cdots + (1+\gamma_{rir})u_r. & \end{array}\right.$$

And this matrix is exactly the l-matrix $E+N_i$ of G_N. Since G_R is generated by the r elements u_e+u_i, $i=1, 2, \cdots, r$, it follows that the group $G_R(\ )$ is simply isomorphic with G_N, and consequently with G (of §4).

This representation of G as a multiplicative group of elements in a finite ring which are congruent to the principal unit modulo a nil ring has conspicuous advantages for the study of the structure of G.† For instance, if $\mathfrak{R}'$ is an invariant subring of $\mathfrak{R}$, then the elements of G_R which have the form

* A nilpotent ring of index k is a ring in which the product of any k elements is 0; a nil ring is one in which some power of every element is 0. A nil ring containing a finite basis is necessarily nilpotent.

† This method of "Strahlbildung" has been used to advantage by K. Shoda (Mathematische Annalen, vol. 100 (1928), pp. 674–686) for determining two important series of characteristic subgroups of $GL_p(\delta_1, \delta_2, \cdots, \delta_r)$.

$[1]u_e+\rho$, ρ in $\mathfrak{R}'$, constitute an invariant subgroup of G_R. Thus the subgroup $\mho_\alpha(G_R)$ corresponds to the subring $\sum_{i=1}^{r}[x_ip^\alpha]u_i$ of $\mathfrak{R}$.

From the representation G_R we can derive a second set of conditions on the coordinates of those matrices in $L_p(\delta_1, \delta_2, \cdots, \delta_r)$ which define automorphisms of G. Now any automorphism of the additive group in $\mathfrak{R}$ is defined by a linear transformation

$$\Theta: \begin{cases} u_1' = y_{11}u_1 + \cdots + y_{1r}u_r \\ \cdots\cdots\cdots\cdots \\ u_r' = y_{r1}u_1 + \cdots + y_{rr}u_r, \end{cases}$$

in which the y_{ij} are least positive residues satisfying the conditions

(11) $$|\, y_{ij}\,| \not\equiv 0 \bmod p;\ y_{ij} \equiv 0 \bmod p^{\delta_j-\delta_i} \text{ for } i > j.$$

But Θ defines a correspondence of the elements of G_R, and for this correspondence to define an automorphism of G_R, it is clearly necessary and sufficient that the multiplication table of the u_i''s be the same as that of the u_i's. By substituting the u_i''s in equations (10) and taking account of the relations $u_i^2\equiv 0$ and $u_ju_i\equiv -u_iu_j$ we find that Θ will define an automorphism of G_R if, and only if, the y's satisfy the $r^2(r-1)/2$ congruences

$$\sum_{\mu,v} D_{ij-\mu v}\gamma_{\mu v k} \equiv \sum_{\alpha=1}^{r} \gamma_{ij\alpha}y_{\alpha k} \bmod p^{\delta_k},$$

where $k = 1, \cdots, r$; $i < j$; $i = 1, \cdots, r$; $j = 2, \cdots, r$; $\mu<v$; $\mu=1, \cdots, r$; $v=2, \cdots, r$, and where $D_{ij-\mu v}$ denotes the determinant

$$\begin{vmatrix} y_{i\mu} & y_{iv} \\ y_{j\mu} & y_{jv} \end{vmatrix}.$$

As one would expect, these congruences are equivalent to the matrix-equations (7) in §5.

The representation G_R above was derived (although indirectly) from a particular ω-normal U-basis of G, and our method is probably the simplest one for obtaining a ring-representation for G. But the existence of a U-basis is by no means necessary* for the existence of a representation of a finite p-group $\overline{G}$ as the multiplicative group of those elements in a ring which are congruent to the principal unit modulo a subring whose elements are expressible uniquely in the form $\sum_{i=1}^{\bar{r}}\xi_iw_i$, the ξ_i being least positive residues

* The quaternion group, which has no U-basis, is the multiplicative group of elements $\rho+\xi_1u_1+\xi_2u_2+\xi_3u_3$, $i=0, 1$, in the ring of 16 elements whose multiplication is defined by the equations $e^2=e$; $eu_i=u_ie=u_i$; $u_1^2=u_2^2=u_2u_1=u_3$; $u_1u_3=u_3u_1=u_1u_2=u_2u_3=u_3u_2=u_3^2=0$.

modulo certain powers of p. Nor is it necessary, when $\overline{G}$ is the group G above, that $\bar{r}$, the number of linearly independent basis units of $\overline{\mathfrak{R}}$, be equal to r, the number of type-invariants of G. For $\bar{r}\neq r$, however, the ring $\overline{\mathfrak{R}}$ must contain elements whose square is not zero, and although it is always nilpotent, it need not be of index 3.

In conclusion, we point out that if the group G is a direct product of groups $G'\times G''\times\cdots$, then for each factor $G^{(i)}$ we can find an ω-normal U-basis, and by the method given in §§4 and 6 we can construct for each factor a ring-representation $G_{R^{(i)}}$. Then for G itself we obtain a ring-representation if we replace $\mathfrak{R}$ in §6 by the direct sum of the rings $\mathfrak{R}'+\mathfrak{R}''+\cdots$. And from this representation we obtain, by post-multiplication, a representation of G as the direct sum of matrix-representations $G_N'+G_N''+\cdots$.

7. In the preceding section we proved that every metabelian group of prime-power order can be exhibited as a multiplicative group in a finite ring. And since every metabelian group is the direct product of its Sylow subgroups, one may construct for any metabelian group $\overline{G}$ of odd order a representation of this sort in which the ring $\mathfrak{R}$ is replaced by the direct sum of nil rings $\mathfrak{R}_{p_1}, \mathfrak{R}_{p_2}, \cdots$, each $\mathfrak{R}_{p_\lambda}$ corresponding to a Sylow subgroup of $\overline{G}$. We wish to show that there is, in a crude sense, a reciprocal relationship between nil rings and metabelian groups.

If S is a ring which contains a principal unit e and a subring Σ such that
(a) *the square of every element in Σ is the zero element, and*
(b) *the number of elements in Σ is an odd integer n,*
then those elements in S which are of the form $e+\sigma$, σ in Σ, constitute under multiplication a group of order n whose class does not exceed 2.

It is easy to show that the elements $e+\sigma$ constitute under multiplication a group F_Σ of order n, having e as the identical operation. We therefore give only the proof that F_Σ is either abelian or metabelian.

Let α and β denote any two elements of Σ. From (a) we have

$$\sigma^2 = 0, \text{ where } \sigma \text{ is any element of } \Sigma. \tag{1}$$

By substituting $\alpha+\beta$ for σ in this equation, we obtain

$$\alpha\beta + \beta\alpha = 0. \tag{2}$$

Two cases arise:

Case A. $\alpha\beta=\beta\alpha$ for every pair of elements in Σ;

Case B. Σ contains two elements σ_α and σ_β such that $\sigma_\alpha\sigma_\beta\neq\sigma_\beta\sigma_\alpha$.

In Case A equation (2) reduces to $2\alpha\beta=0$. Since Σ contains a finite number of elements, with each element σ_i there is associated a smallest positive

integer m_i such that $m_i\sigma_i=0$. Now the elements of Σ constitute under addition an abelian group of order n, and since m_i is clearly the order of σ_i with respect to this group, we see that m_i, being a divisor of n, is necessarily odd. Hence the equation $2\alpha\beta=0$ is possible only if $\alpha\beta=0$. In Case A, therefore, any two elements $e+\alpha$ and $e+\beta$ are commutative, and F_Σ is of class 1.

For Case B we first prove that the product of any three elements of Σ is zero. By using (2) and the associativity postulate, we obtain the equations

$$\begin{aligned}(\sigma_\alpha\sigma_\beta)\sigma_\gamma &= \sigma_\alpha(\sigma_\beta\sigma_\gamma) = -(\sigma_\beta\sigma_\gamma)\sigma_\alpha = -\sigma_\beta(\sigma_\gamma\sigma_\alpha) = (\sigma_\gamma\sigma_\alpha)\sigma_\beta \\ &= \sigma_\gamma(\sigma_\alpha\sigma_\beta) = -(\sigma_\alpha\sigma_\beta)\sigma_\gamma.\end{aligned} \tag{3}$$

That is, $2\sigma_\alpha\sigma_\beta\sigma_\gamma=0$; and as in Case A, we infer that $\sigma_\alpha\sigma_\beta\sigma_\gamma$ is zero. To prove that F_Σ is of class 2 it is sufficient to show that the commutator $(e+\sigma_\alpha, e+\sigma_\beta)$ of any two elements in F_Σ is commutative with any third element $e+\sigma_\gamma$. From (1) we find that the inverse of $e+\sigma_\alpha$ is $e-\sigma_\alpha$. By making use of (3), it is a simple matter to show that the commutator $(e+\sigma_\alpha,\ e+\sigma_\beta)$ is given by $e+\sigma_\alpha\sigma_\beta-\sigma_\beta\sigma_\alpha$ and is commutative with $e+\sigma_\gamma$.

Finally, we observe that the theorem above is valid if we replace (a) by the assumption that Σ is nilpotent and of index 3; that is, the product of any three elements in Σ is the zero-element. (In this case, the commutator $(e+\sigma_\alpha, e+\sigma_\beta)$ equals $e+\sigma_\alpha\sigma_\beta-\sigma_\beta\sigma_\alpha-\sigma_\alpha^2-\sigma_\beta^2$.)

Brown University,
Providence, R. I.

GEOMETRY OF DYNAMICAL TRAJECTORIES AT A POINT OF EQUILIBRIUM*

BY

EDWARD KASNER AND AARON FIALKOW†

1. **Introduction.** A plane positional field of force gives rise to a three-parameter family of curves. These are the totality of possible paths or trajectories of a particle moving under the influence of the field. This system of curves is not arbitrary but has intrinsic peculiarities. Kasner's‡ investigations of the geometry of this situation showed that many aspects of dynamics may be described in geometric language. Excluding points where the force vanishes, he obtained a set of five properties which completely characterize the system of trajectories. Further observations on these five properties are given in recent papers by Moisseiev§ together with new kinematic properties.

In this note, we derive some similar geometric properties of the paths along which a particle may move when it is projected from an isolated point of equilibrium. The new properties are stated in terms of osculating cubics and higher algebraic curves. Since such a point is a singular point of the lines of force, the geometric character of the field is quite different from that at regular points. The qualitative nature of the lines of force follows, of course, from the classic work of Poincaré,‖ Bendixson,¶ and others on the singular points of ordinary differential equations of the first order. A few elementary connections between the differential geometry of the paths and the general character of the lines of force are given in this paper.

2. **The equation of the trajectory.** In what follows we assume that the force is analytic in the neighborhood of the point of equilibrium which we choose as the origin of coordinates. The components of the force are $\phi(x, y)$ and $\psi(x, y)$, where

$$(1)\quad \begin{aligned} \phi(x, y) &= Ax + By + G_{20}x^2 + G_{11}xy + G_{02}y^2 + \cdots, \\ \psi(x, y) &= Cx + Dy + H_{20}x^2 + H_{11}xy + H_{02}y^2 + \cdots. \end{aligned}$$

A particle of unit mass projected from the origin travels along a path whose parametric equations, in terms of the time t, are

* Presented to the Society, October 31, 1936; received by the editors October 17, 1936.

† National Research Fellow.

‡ *Differential-Geometric Aspects of Dynamics*, American Mathematical Society Colloquium Publications, vol. 3, part 2, 1913; new edition, 1934. Also these Transactions, vol. 7 (1906), pp. 401–424.

§ Proceedings of the Sternberg Astronomical Observatory, vol. 7 (1936).

‖ Journal de Mathématiques (1881, 1882); also other papers collected in *Oeuvres*, vol. 1, 1928.

¶ Acta Mathematica, vol. 24 (1901), pp. 1–88.

$$x = a_1t + a_2t^2 + a_3t^3 + \cdots ,$$
$$y = b_1t + b_2t^2 + b_3t^3 + \cdots . \tag{2}$$

These equations are a solution of

$$\frac{d^2x}{dt^2} = \phi(x, y),$$
$$\frac{d^2y}{dt^2} = \psi(x, y). \tag{3}$$

To determine the a_i and b_i, we substitute the values for x and y given by (2) in (1) and (3) and equate like powers of t. The first few equations resulting are

$$\begin{aligned} a_2 &= 0, \\ 3\cdot 2\cdot a_3 &= Aa_1 + Bb_1, \\ 4\cdot 3\cdot a_4 &= G_{20}a_1^2 + G_{11}a_1b_1 + G_{02}b_1^2 , \\ 5\cdot 4\cdot a_5 &= Aa_3 + Bb_3 + G_{30}a_1^3 + G_{21}a_1^2b_1 + G_{12}a_1b_1^2 + G_{03}b_1^3 . \end{aligned} \tag{4}$$

Similar expressions obtain for the b_i. Since a_1 and b_1 are the components of the initial velocity, they are arbitrary parameters. For convenience we write* $m = b_1/a_1$ and $c_i = b_ia_1 - a_ib_1$. If in $\phi(x, y)$ and $\psi(x, y)$, the terms of the ith degree are $X_i(x, y)$ and $Y_i(x, y)$ respectively, we denote by $K_{i+1}(m)$ the polynomial $Y_i(1, m) - mX_i(1, m)$ of degree $(i+1)$ in m. From (4),

$$\begin{aligned} a_2 = b_2 &= 0, \qquad c_2 = 0, \\ 3\cdot 2\cdot c_3 &= a_1^2K_2(m), \\ 4\cdot 3\cdot c_4 &= a_1^3K_3(m), \\ 5\cdot 4\cdot c_5 &= a_1^4K_4(m) + \frac{(A + D)}{3\cdot 2}a_1^2K_2(m). \end{aligned} \tag{5}$$

If we eliminate t between the two equations of (2), the cartesian equation of the path is obtained.† It is

---

* We assume that neither a_1 nor b_1 is zero in the derivation of (7) and (8). If $b_1=0$, it can be shown that the equations of the paths are (7) or (8) with $m=0$. If $a_1=0$, an analogous statement is true if we interchange x and y as well as ϕ and ψ.

† In Kasner's Transactions (1906) paper, it is shown that all the paths are the solutions of the third-order differential equation

$$\left(\psi - \frac{dy}{dx}\cdot\phi\right)\frac{d^3y}{dx^3} = \left\{\psi_x + (\psi_y - \phi_x)\frac{dy}{dx} - \phi_y\left(\frac{dy}{dx}\right)^2\right\}\frac{d^2y}{dx^2} - 3\phi\left(\frac{d^2y}{dx^2}\right)^2.$$

The coefficient of d^3y/dx^3 vanishes at a point of equilibrium. The equations (7) and (8) could also be obtained from this equation.

$$\frac{y}{b_1} = \frac{x}{a_1} + \frac{c_3}{a_1 b_1}\left(\frac{x}{a_1}\right)^3 + \frac{c_4}{a_1 b_1}\left(\frac{x}{a_1}\right)^4 + \frac{a_1 c_5 - 3a_3 c_3}{a_1^2 b_1}\left(\frac{x}{a_1}\right)^5 + \cdots . \tag{6}$$

The values of a_i, b_i, and c_i given by (4) and (5) are substituted in (6). The equation of the path becomes

$$\begin{aligned} y = mx &+ \frac{K_2(m)}{3\cdot 2\cdot a_1^2}x^3 + \frac{K_3(m)}{4\cdot 3\cdot a_1^2}x^4 \\ &+ \frac{1}{5\cdot 4\cdot a_1^2}\left(K_4(m) + \frac{(D - 9A - 10Bm)K_2(m)}{3\cdot 2\cdot a_1^2}\right)x^5 + \cdots . \end{aligned} \tag{7}$$

If $K_2(m)$, $K_3(m)$, $\cdots$, $K_{n-2}(m)$ vanish identically the coefficients of x^3, x^4, $\cdots$, x^{n-1} are zero. The same is true if these polynomials are zero for a particular value of m. In both of these cases, after a complicated discussion we find that the equation of any trajectory having an initial slope m for which $K_2(m)$, $K_3(m)$, $\cdots$, $K_{n-2}(m)$ but not $K_{n-1}(m)$ vanish, is

$$\begin{aligned} y = mx &+ \frac{K_{n-1}(m)}{n(n-1)a_1^2}x^n + \frac{K_n(m)}{n(n+1)a_1^2}x^{n+1} \\ &+ \frac{1}{(n+1)(n+2)a_1^2} \\ &\left(K_{n+1}(m) + \frac{(D - n^2A - (1+n^2)Bm)K_{n-1}(m)}{n(n-1)a_1^2}\right)x^{n+2} + \cdots . \end{aligned} \tag{8}$$

From (7) and (8) it follows that the initial departures from their common tangent of all trajectories projected in the same direction vary inversely as the squares of the initial speeds. In this respect the behavior of the paths is the same at points of equilibrium as it is at regular points. We introduce the intrinsic measure J, defined as the first non-vanishing derivative of the curvature with respect to arc.* We thus have

THEOREM I. *If particles are projected in the same direction from a point in a field of force, the quantities J defined for their paths always vary inversely as the squares of the initial speeds.*

Bendixson† studied the qualitative nature of the lines of force near an

* In previous papers, Kasner used the symbol J (which he termed "inflexure") for the value of the derivative of the curvature of a curve with respect to its arc length at a point of inflexion. It is the invariant measure of the curve at its inflexions analogous to curvature at ordinary points. In this paper, we use this same symbol for the first non-vanishing derivative of the curvature with respect to arc length. Thus J is defined for all analytic curves except straight lines.

† For a proof of his theorems which are used below as well as a precise statement of the conditions under which they are true, see Bendixson, loc. cit., pp. 34, 36, and 62.

isolated singular point of the field. In this paper we are considering the differential-geometric character of the paths. In what follows we indicate certain connections between these two aspects of the field of force.

It was shown by Bendixson that if $X_i(x, y) \equiv Y_i(x, y) \equiv 0$, $(i=1, 2, \cdots, n-3)$, then only the real roots of $K_{n-1}(m)=0$ may be and, in general, are the initial directions of the lines of force which approach the point of equilibrium. If $K_{n-1}(m)$ is identically zero, then every ray from the point (except, possibly, for a finite number of singular rays) is the initial tangent of one and only one line of force. If $K_{n-1}(m)=0$ has no real roots, he showed that the lines of force are either spirals or closed curves about the point. These results of Bendixson together with equation (8) give

THEOREM II. *If particles are projected in different directions from a point of equilibrium, then, in general, any path which is initially tangent to any line of force has higher order of contact with its tangent than a path which is not tangent to a line of force.*

THEOREM III. *If all trajectories obtained by projection from a point of equilibrium have the same order of contact with their tangents, then either there exists a line of force which approaches the point of equilibrium along each ray (except, possibly, for a finite number of rays), or every nearby line of force is either a spiral or a closed curve about the point.*

3. **Geometric properties for the simplest case.** We proceed to translate equations (7) and (8) into geometric language. As is obvious from (7), every trajectory through the origin has at least second order contact with its tangent. At a regular point, on the contrary, the path has a point of inflexion only if the particle is projected in the direction of the force.*

In the simplest case, $X_1(x, y)$, $Y_1(x, y)$, and $K_2(m)$ do not vanish identically. Then, at the origin, from (7),

$$\frac{d^2y}{dx^2} = 0, \tag{9}$$

$$\frac{d^4y}{dx^4} = \frac{2K_3(m)}{K_2(m)} \cdot \frac{d^3y}{dx^3}. \tag{10}$$

According to (9), a curve which osculates a trajectory at the origin must have an inflexion there. We cannot therefore use parabolas as in the ordinary case. The simplest curves with inflexions are cubic curves. We shall restate (9) and (10) as theorems about certain cubic curves which have fourth order contact with the trajectories. The curves that we choose are the cubics which

* Kasner, Proceedings of the National Academy of Sciences, vol. 20 (1934), p. 131, Theorem 3.

have a double point at infinity at which the line at infinity is one of the tangents. Of course, the other tangent is the asymptote of the cubic. We call each of these curves a "special cubic."

In order to simplify the calculations which follow, we first suppose $m=0$. When we take account of (9), the special cubic becomes

$$(y - ax)^3 + by(y - ax) + cy = 0,$$

where a is the slope of the asymptote and c is a non-zero constant.* By a rotation of axes through an angle θ, where $\tan\theta = m$, we obtain the general equation of the special cubic as

$$\begin{aligned}&\cos^3\theta[(1 - am)y - (a + m)x]^3\\ &\qquad + b\cos^2\theta[m(m + a)x^2 + (am^2 - 2m - a)xy + (1 - am)y^2]\\ &\qquad + c\cdot\cos\theta\cdot[y - mx] = 0.\end{aligned}\tag{11}$$

The asymptote of (11) is easily found to be

$$y(1 - am) = (a + m)x - c(1 + m^2)^{1/2}/b.\tag{12}$$

From (11), by differentiation

$$\begin{aligned}\frac{d^3y}{dx^3} &= \frac{6a^3(1 + m^2)^2}{c},\\ \frac{d^4y}{dx^4} &= \frac{4ab(1 + m^2)^{1/2}}{c}\cdot\frac{d^3y}{dx^3}\end{aligned}\tag{13}$$

at the origin. From the equations (10) and (13), it follows that $b/c = K_3(m)/[2aK_2(m)(1+m^2)^{1/2}]$. Substituting this value in (12), the equation of the asymptote becomes

$$y - \frac{2mK_2(m)}{K_3(m)} = \tan(\theta + \alpha)\left(x - \frac{2K_2(m)}{K_3(m)}\right),\tag{14}$$

where $\alpha = \text{arc}\tan a$ and is the angle between the initial tangent and the asymptote. It follows from (14) that if m is constant the asymptotes all intersect at a fixed point which is on the initial tangent, thus forming a pencil. This proves

THEOREM IV. *The special cubics which, at a point of equilibrium, osculate the ∞^1 trajectories passing through the given point in the same direction have asymptotes which form a pencil whose center lies on the common initial tangent.*

* The additional condition that c is a non-zero constant may easily be stated in geometric language. In the general case which is discussed in §4 below, similar restrictions are made.

By reversing the argument given above it is easy to derive a converse statement. We consider any two-parameter family of curves, ∞^1 in each direction, each having an inflexion at the origin. If the asymptotes of the special cubics osculating those curves which have the same initial slope form a pencil with center on the common tangent, then the differential elements at the origin obey the equation

$$\frac{d^4y}{dx^4} = f(m)\cdot\frac{d^3y}{dx^3}. \tag{15}$$

As m varies, the center of the pencil describes a locus. The parametric equations of this locus, by (14), are

$$x = \frac{2K_2(m)}{K_3(m)}, \qquad y = \frac{2mK_2(m)}{K_3(m)}.$$

Eliminating m, we find the cartesian equation of the locus. It is

$$xK_3\left(\frac{y}{x}\right) - 2K_2\left(\frac{y}{x}\right) = 0.$$

This is a cubic curve with a double point at the origin. This equation in conjunction with a theorem of Bendixson previously cited proves

THEOREM V. *The locus of the centers of the ∞^1 pencils of asymptotes corresponding to a point of equilibrium is a cubic of which that point is a double point. The real semi-tangents at the double point are the only rays along which lines of force approach the point of equilibrium. If no real tangents exist, then all the nearby lines of force are either spirals or closed curves about the point.*

Conversely, if the locus of the centers is any cubic curve having a double point at the origin, $f(m)$ in (15) must be a rational function of m whose numerator and denominator are of the third and second degrees respectively.

4. **Geometric properties for the general case.** We now consider the case in which the leading polynomials present in $\phi(x, y)$ and $\psi(x, y)$ are $X_{p-2}(x, y)$ and $Y_{p-2}(x, y)$, $(p \geqq 3)$, respectively.* If $K_{p-1}(m)$ is identically zero there is, in general, one line of force corresponding to each ray from the origin. If $K_{p-1}(m)$ does not vanish identically, the possible rays along which lines of force may approach the origin are given by the real roots of $K_{p-1}(m) = 0$.

In the following we sketch the proofs of analogues of Theorems IV and V for the general case. Let $K_{n-1}(m)$, $(n \geqq p)$, be the first $K_i(m)$ which is not identically zero. Then the paths are given by (8). At the origin

* If $\phi(x, y)$ and $\psi(x, y)$ do not have leading polynomials of the same degree it is always possible to obtain equivalent components of the force of the desired type by a suitable rotation of axes.

$$\frac{d^i y}{dx^i} = 0 \qquad (i = 2, 3, \cdots, n-1), \tag{16}$$

$$\frac{d^{n+1}y}{dx^{n+1}} = \frac{(n-1)K_n(m)}{K_{n-1}(m)} \cdot \frac{d^n y}{dx^n}. \tag{17}$$

As in §3, these equations may be converted into properties of osculating curves. The curves that we select have $(n+2)$ points in common with the trajectories at the origin. They are special n-ics which have a double point at infinity at which the line at infinity has contact of $(n-2)$nd order with the n-ic. The other line having the same contact with the n-ic is its asymptote. Imposing conditions (16), the equation of the special n-ic is (11) after the exponent 3 is replaced by n. From (17) it follows by a method similar to that used in obtaining (14) that the asymptotes of those special n-ics which have the same m form a pencil whose center is $([(n+1)K_{n+1}(m)]/[(n-1)K_n(m)]$, $[(n+1)mK_{n-1}(m)]/[(n-1)K_n(m)])$. We thus have

THEOREM VI. *The special n-ics which, at a point of equilibrium, osculate the ∞^1 trajectories passing through the point in the same direction have asymptotes which form a pencil whose center lies on the initial tangent.*

As m varies, the center of the pencil describes a locus whose equation is

$$(n-1)xK_n\left(\frac{y}{x}\right) - (n+1)K_{n-1}\left(\frac{y}{x}\right) = 0.$$

As in §3, this proves

THEOREM VII. *The locus of the centers of the ∞^1 pencils of asymptotes corresponding to a point of equilibrium is an n-ic of which that point is an $(n-1)$-point. If $n=p$, the real semi-tangents at the $(n-1)$-point are the only rays along which lines of force may approach the point of equilibrium. If no real tangents exist, then all the nearby lines of force are either spirals or closed curves about the point. For $n>p$, there is, in general, one and only one line of force which approaches the point of equilibrium along each ray (except, possibly, for a finite number of rays).*

A reversal of the argument employed to obtain these theorems yields converse statements similar to those of §3.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.
PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

ON CERTAIN INEQUALITIES AND CHARACTERISTIC VALUE PROBLEMS FOR ANALYTIC FUNCTIONS AND FOR FUNCTIONS OF TWO VARIABLES*

BY
KURT FRIEDRICHS

PART I. THE CASE OF ANALYTIC FUNCTIONS

1. INTRODUCTION

In this first part I investigate some properties of the manifold $\mathfrak{F}$ of all analytic functions $u+iv=w(z)$ defined in a bounded open connected domain D of the $(z=x+iy)$-plane for which the integral

$$\iint_D |w|^2 dxdy$$

is finite.†

First I establish the following inequality. There exists a positive constant $\theta<1$ such that, for all functions $w(z)$ which satisfy the additional condition

$$\iint wdxdy = 0,$$

the inequality

$$\left| \iint_D w^2 dxdy \right| \leqq \theta \iint_D |w|^2 dxdy \tag{1.1}$$

is valid.

It will be seen that this inequality is equivalent to

$$\iint_D u^2 dxdy \leqq \Gamma \iint_D v^2 dxdy \tag{1.2}$$

under the additional condition

$$\iint_D udxdy = 0,$$

the constant $\Gamma=(1+\theta)/(1-\theta)$ being greater than 1.

* Presented to the Society, October 31, 1936; received by the editors December 6, 1935, and January 27, 1936.

† The same space was investigated in regard to different properties by St. Bergman, Mathematische Annalen, vol. 86 (1922), p. 238, and Berliner Sitzungsberichte, 1927, p. 178; and S. Bochner, Mathematische Zeitschrift, vol. 14(1922), p. 180).

Secondly, I deal with the characteristic value problem for the quadratic form

$$\iint_D w^2 dxdy$$

with respect to the unit-form

$$\iint_D |w|^2 dxdy.$$

I prove the existence of a sequence of characteristic values $\mu_1, \mu_2, \mu_3, \cdots, \mu_n \downarrow 0$, and corresponding characteristic functions $w_1(z), w_2(z), \cdots$ satisfying the conditions

$$\iint_D \overline{w_m(z)} w_n(z) dxdy = \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases}$$

and thus being orthonormal, such that each function $w(z)$ in $\mathfrak{F}$ can be developed in a series

$$w(z) = c_1 w_1(z) + c_2 w_2(z) + c_3 w_3(z) + \cdots$$

converging uniformly in every closed subdomain of D, while the expansions

$$\iint_D |w|^2 dxdy = |c_1|^2 + |c_2|^2 + |c_3|^2 + \cdots$$

$$\iint_D w^2 dxdy = \mu_1 c_1^2 + \mu_2 c_2^2 + \mu_3 c_3^2 + \cdots$$

hold.

The largest characteristic value μ_1 is equal to 1 and the corresponding characteristic function $w_1(z)$ is constant. The inequality stated above expresses nothing but the fact that the second characteristic value μ_2 is less than 1.

The validity of these theorems depends essentially on the nature of the boundary of the domain D. My assumption is that this boundary B consists of a finite number of closed curves having a continuous tangent except at a finite number of corners. Then the inequality holds; but the expansion theorem is valid if and only if there are no corners (except internal cusps). In the case of corners the extreme points of the limit spectrum can be determined.

Let each closed curve of the boundary be represented by a continuous

periodic function $z=z(s)$ of a parameter s. Except at a finite number of corners this function shall have a continuous derivative

$$\frac{dz}{ds}=\dot{z}(s),$$

which, at each corner, is continuous on both sides. We can assume

$$|\dot{z}| = \left|\frac{dz}{ds}\right| = 1,$$

so that s is the arc length of the curve, and further that the normal $i\dot{z}$ is directed into the interior of D.

At a corner the argument of $\dot{z}$ has the jump $(1-\omega)\pi$, where $\omega\pi$ is the inner angle of the corner.

We assume $0<\omega\leqq 2$. Thus we exclude external cusps $(\omega=0)$; in this case it can happen that not even the inequality holds as we shall show at the end of this part.

2. A basic lemma

We set

$$\mathrm{M} = \max_{\omega}\left|\frac{\sin \omega\pi}{\omega\pi}\right|.$$

Since the case $\omega=0$ has been excluded we have $\mathrm{M}<1$, and $\mathrm{M}=0$ if there are no corners (except internal cusps).

LEMMA 1. *Let ϵ be an arbitrary positive number. Then there exists a boundary strip S in D, bounded by the exterior boundary B and an interior boundary B' which consists of a finite number of rectifiable closed curves, such that for every function $w(z)=u(z)+iv(z)$ in $\mathfrak{F}$ the inequality*

$$\left|\iint_S w^2dxdy\right| \leqq (\mathrm{M}+\epsilon)\iint_S |w|^2dxdy + \gamma\int_{B'}|w|^2|dz| \tag{2.1}$$

holds, γ being a suitable positive constant depending on ϵ. This inequality implies

$$(1-\mathrm{M}-\epsilon)\iint_S u^2dxdy \leqq (1+\mathrm{M}+\epsilon)\iint_S v^2dxdy + \gamma\int_{B'}|w|^2|dz|. \tag{2.2}$$

Without loss of generality we may confine ourselves to the case of a domain D bounded by only one closed curve B.

First we discuss the case of no corners. We may assume $\epsilon<1$. We choose a constant σ in such a way that

$$\left|\frac{z(s')-z(s)}{s'-s} - \dot{z}(s)\right| \leqq \frac{\epsilon}{4} \quad \text{as} \quad |s'-s| \leqq \sigma. \tag{2.3}$$

We set

$$\Delta z(s) = \frac{z(s+\sigma) - z(s)}{\sigma}, \qquad \Delta \dot{z}(s) = \frac{\dot{z}(s+\sigma) - z(s)}{\sigma}.$$

Then we have

$$(2.3)' \qquad |\Delta z(s) - \dot{z}(s)| \leqq \frac{\epsilon}{4},$$

$$\frac{3}{4} \leqq |\Delta z(s)| \leqq \frac{5}{4}.$$

We choose a number $\rho > 0$ in such a way that

$$|s' - s| \leqq \sigma \quad \text{as} \quad |z(s') - z(s)| \leqq \rho$$

and a number $T > 0$ in such a way that

$$T \leqq \frac{2}{5}\rho$$

and

$$(2.4) \qquad T\left|\frac{\Delta z(s') - \Delta z(s)}{s' - s}\right| \leqq \frac{\epsilon}{40} \quad \text{as} \quad |s' - s| \leqq \sigma.$$

Consequently we have

$$(2.4)' \qquad T|\overline{\Delta z}\Delta \dot{z}| \leqq \frac{\epsilon}{32}.$$

Now we introduce a new parameter t, $0 \leqq t \leqq T$, and set

$$z = z(s) + it\Delta z(s).$$

The strip $0 < t < T$ corresponds in a one-to-one way to a certain boundary strip S in D, bounded by the boundary B and an inner curve B' which correspond to $t = 0$ and $t = T$ respectively. To show this we first prove the relation

$$(2.5) \qquad |z' - z - [(s' - s) + i(t' - t)]\dot{z}(s)| \leqq \frac{3}{8}\epsilon|(s' - s) + i(t' - t)|,$$

as $|s' - s| \leqq \sigma$.

In fact, in view of (2.3), (2.4), we have

$$\begin{aligned} &|z' - z - [(s' - s) + i(t' - t)]\dot{z}(s)| \\ &\quad = |[z(s') - z(s) - (s' - s)\dot{z}(s)] + it'[\Delta z(s') - \Delta z(s)] \\ &\quad\quad + i(t' - t)[\Delta z(s) - \dot{z}(s)]| \end{aligned}$$

$$\leqq \frac{\epsilon}{4}|s'-s| + \frac{\epsilon}{40}|s'-s| + \frac{\epsilon}{4}|t'-t|$$

$$\leqq \left|\frac{\epsilon}{4}+\frac{\epsilon}{40}+i\frac{\epsilon}{4}\right| |(s'-s)+i(t'-t)|.$$

Now let $z'=z$. Then we have

$$\begin{aligned} |z(s')-z(s)| &= |t'\Delta z(s') - t\Delta z(s)| \\ &\leqq t'|\Delta z(s')| + t|\Delta z(s)| \\ &\leqq 2T\frac{5}{4} \leqq \rho. \end{aligned}$$

Therefore $|s'-s|\leqq\sigma$ and relation (2.5) leads to $|(s'-s)+i(t'-t)|(1-\frac{3}{8}\epsilon)\leqq 0$. Hence $s'=s$, $t'=t$. Thus the one-to-one correspondence of $0<t<T$ and S is proved.

We calculate the Jacobian

$$J = \frac{\partial(x,y)}{\partial(s,t)} = \Re\overline{\Delta z}(\dot z + it\Delta\dot z);$$

from (2.3)′, (2.4)′ we find

$$J \geqq 1 - \frac{\epsilon}{4} - \frac{\epsilon}{32} \geqq \frac{1}{2}. \tag{2.6}$$

In the strip S the parameters s and t can be expressed in terms of x, y, hence in terms of $z=x+iy$, $\bar z=x-iy$; and a simple calculation yields the relation

$$2J\frac{\partial}{\partial z} = \overline{\Delta z}\frac{\partial}{\partial s} - \overline{i(\dot z + it\Delta\dot z)}\frac{\partial}{\partial t}.$$

We now introduce the function

$$j(z,\bar z) = z(s) - it\Delta z(s)$$

and find

$$J\frac{\partial j}{\partial z} = i\Im(\dot z\overline{\Delta z}) + t\Im(\Delta\dot z\overline{\Delta z}).$$

From (2.3)′, (2.4)′, (2.6) we get

$$\left|\frac{\partial j}{\partial z}\right| \leqq \epsilon.$$

By a simple calculation, we obtain the identity

$$\iint_{\tau<t<T}\left(1-\frac{\partial\bar{j}}{\partial z}\right)w^2dxdy = \iint_{\tau<t<T}\frac{\partial}{\partial\bar{z}}[(\overline{z-j})w^2]dxdy$$
$$= \frac{i}{2}\int_{B'}(\overline{z-j})w^2dz$$
$$- \frac{i}{2}\int_{t=\tau}(\overline{z-j})w^2dz$$

for $0<\tau<T$. Since $\int_0^T\int_{t=\tau}|w|^2dsd\tau \leqq 2\iint_S|w|^2dxdy$ is finite, there exists a sequence $\tau\to 0$ for which $\int_{t=\tau}|w|^2ds$ is bounded. If we let τ tend to zero in this way, we have

$$\left|\frac{i}{2}\int_{t=\tau}(\overline{z-j})w^2dz\right| = \left|\int_{t=\tau}t\overline{\Delta z}w^2(\dot{z}+it\Delta\dot{z})ds\right|$$
$$\leqq \tau\left(1+\frac{\epsilon}{2}\right)\int_{t=\tau}|w|^2ds \to 0$$

and therefore the identity

$$\iint_S\left(1-\frac{\partial\bar{j}}{\partial z}\right)w^2dxdy = T\int_{B'}w^2\overline{\Delta z}dz.$$

It yields immediately the inequality

$$\left|\iint_S w^2dxdy\right| \leqq \frac{5}{4}T\int_{B'}|w|^2|dz| + \epsilon\iint_S|w|^2dxdy.$$

3. Case of corners

We now pass on to the case of boundary B having corners z_ν with the angles $\omega_\nu\pi$ ($\nu=1,\cdots,n$). We map the domain D conformally on a domain D^* of the z^*-plane such that the boundary B^* of D^* has a continuous tangent. There exists such a mapping, regular in $D+B$ except at the corners, which behaves at the corners as follows: there is an analytic function

$$\lambda = \lambda(z) \quad \text{with} \quad \lambda(z_\nu) = 0, \quad \lambda'(z_\nu) \neq 0,$$

regular in the neighborhood of $z=z_\nu$, and an analytic function

$$\lambda_* = \lambda_*(z^*) \quad \text{with} \quad \lambda_*(z_\nu^*) = 0, \quad \lambda_*'(z_\nu^*) \neq 0$$

(where z_ν^* corresponds to the corner z_ν), regular in the neighborhood of $z^*=z_\nu^*$, such that

$$\lambda = \lambda_*^{\omega_\nu}.\dagger$$

We shall make use of the fact that there is a constant $\chi>0$ such that

$$(3.1)\qquad \left|(z^* - z_\nu^*)\frac{d}{dz^*}\log\frac{dz}{dz^*} + (1-\omega_\nu)\right| \leqq \chi\left|z^* - z_\nu^*\right| + \chi\left|z^* - z_\nu^*\right|^{\omega_\nu}$$

in a certain neighborhood U_ν^* of z_ν^*. We see this by a simple calculation:

$$\frac{dz^*}{dz} = \frac{\lambda_*^{1-\omega_\nu}}{\omega_\nu}\frac{\lambda'}{\lambda_*'},$$

$$\lambda_*\frac{d}{d\lambda_*}\log\frac{dz^*}{dz} = (1-\omega_\nu) - \lambda_*\frac{\lambda_*''}{(\lambda_*')^2} + \omega_\nu\lambda\frac{\lambda''}{(\lambda')^2}.$$

Since

$$\frac{z^* - z_\nu^*}{\lambda_*} : \frac{dz^*}{d\lambda_*} = 1 + (z^* - z_\nu^*)\vartheta_*(z^*),$$

where $\vartheta_*(z^*)$ is bounded at $z^* = z_\nu^*$, we get

$$(z^* - z_\nu^*)\frac{d}{dz^*}\log\frac{dz^*}{dz} - (1-\omega_\nu) = (1-\omega_\nu)(z^* - z_\nu^*)\vartheta_*(z^*) - (z^* - z_\nu^*)\frac{\lambda_*''}{\lambda_*'}$$

$$+ \omega_\nu(z^* - z_\nu^*)^{\omega_\nu}\left(\frac{z^* - z_\nu^*}{\lambda_*}\right)^{1-\omega_\nu}\frac{\lambda_*'}{\lambda'}\frac{\lambda''}{\lambda'}.$$

Since $\lambda_*/(z^* - z_\nu^*) \to \lambda_*'(0) \neq 0$ as $z^* \to z_\nu^*$ and $\lambda'(0) \neq 0$, there is a number $\chi > 0$ and a neighborhood U_ν^* of z_ν^*, where

$$|1-\omega_\nu|\,|\vartheta_*| + \left|\frac{\lambda_*''}{\lambda_*'}\right| \leqq \chi, \qquad \omega_\nu\left|\frac{z^* - z_\nu^*}{\lambda^*}\right|^{1-\omega_\nu}\left|\frac{\lambda_*'}{\lambda'}\right|\left|\frac{\lambda''}{\lambda'}\right| \leqq \chi$$

and, thus, (3.1) holds in U_ν^*. We introduce the function

$$\eta = \frac{dz}{dz^*} : \overline{\frac{dz}{dz^*}} \quad \text{with} \quad |\eta| = 1.$$

† We can construct such a mapping, e.g., in the following way. We choose a number a_1 in the exterior of $D+B$ such that the function

$$z^{(1)} = \left(\frac{z - z_1}{z - a_1}\right)^{1/\omega_1}$$

maps the domain $D+B$ in a one-to-one way on a domain D^1+B^1 of the $z^{(1)}$-plane; D^1+B^1 has no corner at the point $z^{(1)} = z_1^{(1)}$. In the same way we choose $a_2^{(1)}$ in the exterior of D^1+B^1 and form

$$z^{(2)} = \left(\frac{z^{(1)} - z_2^{(1)}}{z^{(1)} - a_2^{(1)}}\right)^{1/\omega_2},$$

and so on. Then we set

$$z^* = z^{(n)}$$

and take $\lambda_* = z^{(\nu)}$ for the corner $z = z_\nu$.

We observe that relation (3.1) is equivalent to

$$(3.2)\qquad \left| (z^* - z_\nu^*)\frac{\partial \eta}{\partial z^*} + (1-\omega_\nu)\eta \right| \leqq \chi\,| z^* - z_\nu^* | + \chi\,| z^* - z_\nu^* |^{\omega_\nu},$$

where η is considered as a function of z^* and $\bar{z}^*$. On setting

$$w^*(z^*) = w(z)\,\frac{dz}{dz^*}$$

we have

$$\iint_D | w |^2 dxdy = \iint_{D^*} | w^* |^2 dx^* dy^*$$

$$\iint_D w^2 dxdy = \iint_{D^*} (w^*)^2 \bar{\eta}\, dx^* dy^*.$$

4. Continuation

In what follows we omit the sign * and write z, z_ν, D, B, S, U_ν, w instead of z^*, z_ν^*, D^*, B^*, S^*, U_ν^*, w^*.

Let the boundary B be represented by $z = z(s)$, and let s_ν be that value of s for which $z(s_\nu) = z_\nu$; we set $\dot{z}(s_\nu) = \dot{z}_\nu$. We introduce the parameters s and t in the boundary strip as before (§2) and define the function $j(z, \bar{z})$.

$$z = z(s) + it\Delta z(s); \qquad j = z(s) - it\Delta z(s).$$

Let S_ν and B_ν be the domain of all points of S and B respectively for which $|s - s_\nu| \leqq \sigma$, except $z = z_\nu$. We assume $\epsilon < \frac{1}{3}$ and at the same time so small that $B_\nu + S_\nu$ is contained within the neighborhood U_ν of $z = z_\nu$. From (3.2) we get

$$(4.1)\qquad \left| (z - z_\nu)\frac{\partial \eta}{\partial z} + (1-\omega_\nu)\eta \right| \leqq \chi\,| z - z_\nu | + \chi\,| z - z_\nu |^{\omega_\nu} \quad \text{in} \quad S_\nu + B_\nu.$$

We take note of the relations

$$(4.2)\qquad \begin{aligned} &| (z - z_\nu) - (s - s_\nu + it)\dot{z}_\nu | \leqq \frac{3}{8}\epsilon\, | s - s_\nu + it | \\ &| (j - z_\nu) - (s - s_\nu - it)\dot{z}_\nu | \leqq \frac{3}{8}\epsilon\, | s - s_\nu - it | \end{aligned}$$

which holds for $|s - s_\nu| \leqq \sigma$ and, therefore, in $S_\nu + B_\nu$. They can be derived in the same way as relation (2.5).

We define the functions

$$K_\nu(z, \bar{z}) = -\frac{z - z_\nu}{\dot{z}_\nu} : \frac{j - z_\nu}{\dot{z}_\nu}$$

$$\mathrm{H}_\nu(z, \bar{z}) = -\frac{z - z_\nu}{\dot{z}_\nu} : \overline{\frac{z - z_\nu}{\dot{z}_\nu}}$$

in $S_\nu + B$; we have

(i) $|\mathrm{H}_\nu| = 1$,

(ii) $1 - \epsilon \leqq |\mathrm{K}_\nu| \leqq 1 + \epsilon$,

(iii) $|\mathrm{H}_\nu - \mathrm{K}_\nu| \leqq \epsilon$.

In view of (4.2) relation (iii) follows from

$$|\mathrm{H}_\nu - \mathrm{K}_\nu| = \frac{\left| \overline{\dfrac{z - z_\nu}{\dot{z}_\nu}} - \dfrac{j - z_\nu}{\dot{z}_\nu} \right|}{\left| \dfrac{j - z}{\dot{z}_\nu} \right|} \leqq \frac{\dfrac{3}{4}\epsilon}{1 - \dfrac{3}{8}\epsilon} \leqq \epsilon.$$

The relations (iii) and (i) lead to (ii).

The function K_ν has the value 1 on the lines $s = s_\nu$, $0 < t < T$, and the value -1 on the boundary $t = 0$, $s \neq s_\nu$. Therefore $\log \mathrm{K}_\nu$ is defined; we have

$$\begin{aligned} \log \mathrm{K}_\nu &= 0 \quad &&\text{on} \quad s = s_\nu, \quad &&0 < t < T; \\ \log \mathrm{K}_\nu &= \mp i\pi \quad &&\text{on} \quad t = 0, \quad &&s \lessgtr s_\nu. \end{aligned}$$

In view of (4.2) we get

$$|\log \mathrm{K}_\nu| \leqq \frac{\dfrac{3}{4}\epsilon}{1 - \dfrac{3}{8}\epsilon} + \left| \log \frac{t + i(s - s_\nu)}{t - i(s - s_\nu)} \right| \leqq \epsilon + \pi.$$

We can define $\log \mathrm{H}_\nu$ in such a way that $|\log \mathrm{H}_\nu| \leqq 3\epsilon/2$ on the line $s = s_\nu$ and we set

$$\mathrm{K}_\nu^{\omega_\nu} = \exp \omega_\nu \log \mathrm{K}_\nu, \qquad \mathrm{H}_\nu^{\omega_\nu} = \exp \log \mathrm{H}_\nu.$$

From (iii) we obtain

(iv) $|\mathrm{K}_\nu^{\omega_\nu} - \mathrm{H}_\nu^{\omega_\nu}| \leqq 3\epsilon$.

We consider the function

$$\mathrm{P}_\nu(z, \bar{z}) = \frac{1}{\omega_\nu} \mathrm{K}_\nu^{\omega_\nu} - \frac{\cos \omega_\nu \pi}{\omega_\nu} - \frac{\sin \omega_\nu \pi}{\omega_\nu \pi} \log \mathrm{K}_\nu,$$

defined in $S_\nu + B_\nu$; it vanishes on the boundary $t = 0$, $s \neq s_\nu$, and it is bounded. Therefore we have

$$(4.3) \qquad \begin{aligned} (z - z_\nu)\mathrm{P}_\nu(z, \bar{z}) &\to 0 \\ (z - z_\nu)^{\omega_\nu}\mathrm{P}_\nu(z, \bar{z}) &\to 0 \end{aligned} \qquad \text{as } t \to 0$$

uniformly in s. We calculate the derivative of P_ν with respect to z and find

$$(z - z_\nu)\frac{\partial P_\nu}{\partial z} = \left[K_\nu^{\omega_\nu} - \frac{\sin \omega_\nu \pi}{\omega_\nu \pi}\right]\left(1 + K_\nu \frac{\partial j}{\partial z}\right).$$

According to (ii), (iv) and

$$\left| \frac{\partial j}{\partial z} \right| \leqq \epsilon \tag{4.4}$$

we get

$$\left| (z - z_\nu)\frac{\partial P_\nu}{\partial z} - H_\nu^{\omega_\nu} \right| \leqq \left| \frac{\sin \omega_\nu \pi}{\omega_\nu \pi} \right| + 6\epsilon. \tag{4.5}$$

We choose non-negative functions $\rho_\nu(s)$ which have continuous derivatives, which vanish outside of B_ν, which are equal to 1 for $|s - s_\nu| \leqq \sigma/2$ and for which

$$\sum_\nu \rho_\nu(s) \leqq 1.$$

Here the summation $\sum_\nu$ is extended over all points $z = z_\nu$. We consider the function

$$\Omega(z, \bar{z}) = \sum_\nu \rho_\nu(s) H_\nu^{-\omega_\nu}(z - z_\nu)\eta P_\nu + \left[1 - \sum_\nu \rho_\nu(s)\right]\eta(z - j).$$

Since

$$\rho_\nu(s)(z - z_\nu)P_\nu \to 0, \qquad z - j \to 0 \quad \text{as} \quad t \to 0, \tag{4.6}$$

we have

$$\Omega(z, \bar{z}) \to 0 \quad \text{as} \quad t \to 0.$$

We investigate the derivative $\partial\Omega/\partial z$. First we observe that $\partial\rho_\nu/\partial z = \rho_\nu'(\partial s/\partial z)$ is bounded in S and that $\partial\eta/\partial z$ is bounded in $S - \sum_\nu S_\nu$. Further we calculate

$$\frac{\partial}{\partial z} H_\nu^{-\omega_\nu}(z - z_\nu)\eta = (1 - \omega_\nu)H_\nu^{-\omega_\nu}\eta + H_\nu^{-\omega_\nu}(z - z_\nu)\frac{\partial\eta}{\partial z}.$$

We now use relation (4.1); because of $|H_\nu| = 1$ we get

$$\left| \frac{\partial}{\partial z} H_\nu^{-\omega_\nu}(z - z_\nu)\eta \right| \leqq \chi |z - z_\nu| + \chi |z - z_\nu|^{\omega_\nu}$$

in S_ν. Thus, on account of (4.3) and (4.6) we find

$$\sum_\nu P_\nu \frac{\partial}{\partial z}\rho_\nu H_\nu^{-\omega_\nu}(z - z_\nu)\eta + (z - j)\frac{\partial}{\partial z}\left[1 - \sum_\nu \rho_\nu\right]\eta \to 0 \tag{4.7}$$

as $t \to 0$, uniformly in s.

Consequently we can choose a positive number $T_1 \leqq T$ such that for $0 < t \leqq T_1$

$$\left| \frac{\partial \Omega}{\partial z} - \sum_\nu \rho_\nu H_\nu^{-\omega_\nu}(z - z_\nu)\eta \frac{\partial P_\nu}{\partial z} - \left[1 - \sum_\nu \rho_\nu\right]\eta\left(1 - \frac{\partial j}{\partial z}\right)\right| \leqq \epsilon.$$

According to (4.4) and (4.5) this relation leads to

$$\left| \frac{\partial \Omega}{\partial z} - \eta \right| \leqq \sum_\nu \rho_\nu \left| \frac{\sin \omega_\nu \pi}{\omega_\nu \pi} \right| + 7\epsilon \leqq M + \epsilon_1, \tag{4.8}$$

where $\epsilon_1 = 7\epsilon$.

Let S_1 be the strip $0 < t < T_1$ and B_1' the curve $t = T_1$; since $\Omega \to 0$ as $t \to 0$ we have the identity

$$\iint_{S_1} w^2 \frac{\partial \bar{\Omega}}{\partial z} dxdy = \frac{i}{2} \int_{B_1'} w^2 \bar{\Omega} dz.$$

From this and relation (4.8) we deduce

$$\iint_{S_1} w^2 \bar{\eta} dxdy \leqq \gamma_1 \int_{B_1'} |w|^2 |dz| + (M + \epsilon_1) \iint_{S_1} |w|^2 dxdy$$

where $\gamma_1 = 2 \max |\Omega|$ on B_1'. Since ϵ_1 is arbitrarily small we thus have proved Lemma 1 also for the case where the boundary has corners.

5. Additional lemmas

In proving our theorems we further make use of the following elementary lemmas.

Lemma 2. *Let D' be a closed domain within D. Then for all functions $w(z)$ in $\mathfrak{F}$ and all z' in D' we have*

$$|w(z')|^2 \leqq C_0' \iint_D |w|^2 dxdy \tag{5.1}$$

$$\left| \frac{dw}{dz}(z') \right|^2 \leqq C_1' \iint_D |w|^2 dxdy, \tag{5.2}$$

C_0', C_1' being positive constants depending on D'.

We omit the proof of this well known lemma, which is an immediate consequence of the mean-value theorem and the Schwarz inequality.

Lemma 3. *Let D' be a closed subdomain within D and z_0 a point of D'. Then for all functions $w(z) = u(z) + iv(z)$ for which*

$$u(z_0) = 0$$

the inequality

$$|\, w(z')\,|^2 \leqq C' \iint_D v^2 dxdy, \qquad z' \text{ in } D',$$

is valid, C' being a positive constant depending on D' and z_0.

We choose a finite number of points $z_0, z_1, z_2, \cdots, z_j, \cdots, z_k$ in D' and a positive number R such that $|z_j - z_{j-1}| < R$, that every point z' of D' belongs to one of the circles $|z - z_j| < R$ and that the circles $|z - z_j| < 3R$ are in the interior of D. Now we take the relation

$$w(z') - u(z_j) = \frac{1}{2\pi} \int_{|z-z_j|=r} v(z) \frac{z + z' - 2z_j}{z - z'} \frac{dz}{z - z_j}$$

which holds for $|z' - z_j| < r$. We multiply by rdr and integrate with respect to r from $r = 2R$ to $r = 3R$ and apply the Schwarz inequality assuming $|z' - z_j| < R$. Thus we obtain

$$|\, w(z') - u(z_j)\,| \leqq 4\left(\frac{1}{5R^2\pi} \iint_{2R<|z-z_j|<3R} v^2 dxdy\right)^{1/2} \leqq \frac{2}{R}\left(\iint_D v^2 dxdy\right)^{1/2}.$$

Since $u(z_0) = 0$ and $|z_j - z_{j-1}| < R$, we get for every point z' of D'

$$|\, w(z')\,| \leqq (k + 1)\, \frac{2}{R}\left(\iint_D v^2 dxdy\right)^{1/2}.$$

Herewith we have proved Lemma 3.

6. Fundamental inequalities

We now are ready for the proof of

Theorem 1. *There exists a positive number Γ such that the inequality*

$$\iint_D u^2 dxdy \leqq \Gamma \iint_D v^2 dxdy$$

holds for all functions $w = u + iv$ in $\mathfrak{F}$ which satisfy the condition

$$\iint_D u dxdy = 0.$$

We first observe that it is sufficient to prove the inequality for all functions $w = u + iv$ in $\mathfrak{F}$ which vanish at a certain point z_0 in D; this fact follows from the relation

$$\iint_D u^2(z)dxdy \leqq \iint_D [u(z) - u(z_0)]^2 dxdy,$$

which is an immediate consequence of $\iint_D udxdy = 0$. Now we choose a positive number ϵ such that $\mathrm{M}+\epsilon<1$; this is possible since $\mathrm{M}<1$. By (2.2) of Lemma 1 we have

$$(1 - \mathrm{M} - \epsilon)\iint_D u^2 dxdy \leqq \iint_{D-S} u^2 dxdy + (1 + \mathrm{M} + \epsilon)\iint_S v^2 dxdy$$
$$+ \gamma \int_{B'} | w |^2 | dz |.$$

We choose a point z_0 of $D-S$ and assume $u(z_0)=0$. Then we apply Lemma 3 to the closed domain $D'=D-S$ and get

$$(1 - \mathrm{M} - \epsilon)\iint_D u^2 dxdy$$
$$\leqq \left[C' \iint_{D-S} dxdy + (2 + \epsilon) + C' \int_{B'} | dz |\right] \iint_D v^2 dxdy;$$

thus Theorem 1 is proved.

The inequality of Theorem 1 is equivalent to

$$(\Gamma + 1)\Re \iint_D w^2 dxdy \leqq (\Gamma - 1)\iint_D | w |^2 dxdy,$$

and the additional condition is the same as

$$\Re \iint wdxdy = 0.$$

Now let w be a function in $\mathfrak{F}$ which satisfies the relation

$$\iint_D wdxdy = 0;$$

then we can choose a number η of absolute value 1 such that

$$\Re \iint_D (\eta w)^2 dxdy = \left| \iint_D w^2 dxdy \right| ;$$

since $\Re\iint_D \eta wdxdy=0$ we can apply Theorem 1 to ηw. On setting $\theta=(\Gamma-1)/(\Gamma+1)$ we thus obtain

THEOREM 2. *There exists a constant $\theta<1$ such that the inequality*

$$\left| \iint_D w^2 dxdy \right| \leqq \theta \iint_D |w|^2 dxdy$$

holds for all functions in $\mathfrak{F}$ satisfying the condition

$$\iint_D w dxdy = 0.$$

7. SPACE $\mathfrak{F}$ AS A HILBERT SPACE

To prove our expansion theorem we start with the observation that the manifold $\mathfrak{F}$ of all functions $w(z)$ for which

$$\iint_D |w|^2 dxdy$$

is finite constitutes a linear metric space. This space is either complex or real depending on whether we allow multiplication by complex or real numbers and define an inner product either by

$$(w_1, w_2) = \iint_D \bar{w}_1 w_2 dxdy$$

or by

$$(w_1, w_2) = \Re \iint_D \bar{w}_1 w_2 dxdy.$$

In either case the modulus

$$(w, w)^{1/2} = \left(\iint_D |w|^2 dxdy \right)^{1/2}$$

has the same value. In order that

$$\Re \iint_D w_1 w_2 dxdy$$

become a symmetric bilinear form we choose the second definition of (w_1, w_2) and thus assume $\mathfrak{F}$ to be a real space. In this case, iw belongs to the space $\mathfrak{F}$ together with w_1; but this function iw is orthogonal to w:

$$(w, iw) = \Re \iint_D \bar{w} iw dxdy = \Re i \iint_D |w|^2 dxdy = 0.$$

Now we establish

LEMMA 4. *The space $\mathfrak{F}$ of functions $w(z)$ is a Hilbert space.*

First we prove that $\mathfrak{F}$ is complete. Let $w^n(z)$ be a sequence in $\mathfrak{F}$ such that

$$(w^m - w^n, w^m - w^n) = \iint_D | w^m - w^n |^2 dxdy \to 0 \quad \text{as} \quad m, n \to \infty .$$

From the inequality (1) of Lemma 2 we deduce that $w^m(z) - w^n(z)$ converges to zero uniformly in every closed subdomain D' of D. From this fact it follows by a well known procedure that (i) $\iint_D |w|^2 dxdy < \infty$ and therefore $w(z)$ belongs to the space $\mathfrak{F}$; and (ii)

$$\iint_D | w^n - w |^2 dxdy \to 0 \quad \text{as} \quad n \to \infty .$$

$\mathfrak{F}$ is a subspace of the linear space† of all complex-valued functions $k(x, y)$, continuous in D, for which

$$(k \mid k) = \iint_D | k |^2 dxdy < \infty ;$$

since, obviously, this space is separable $\mathfrak{F}$ has a like property. Therefore $\mathfrak{F}$ is a Hilbert space.

A sequence of functions $w^1(z), w^2(z), \cdots, w^n(z), \cdots$ in $\mathfrak{F}$ is called *weakly convergent* to a function $w(z)$ in $\mathfrak{F}$,

$$w^n(z) \overset{\cdot}{\to} w(z) ,$$

if it has the following two properties:

(i) there is a constant $M_0 > 0$ such that

$$\iint_D | w^n |^2 dxdy \leqq M_0;$$

(ii) $$\iint_D \overline{W}(w^n - w) dxdy \to 0 \quad \text{as} \quad n \to \infty$$

for each function $W(z)$ in $\mathfrak{F}$.

We have the

LEMMA 5. *Let $w^1(z), w^2(z), \cdots, w^n(z)$ be a sequence of functions in $\mathfrak{F}$ for which*

$$\iint_D | w^n |^2 dxdy \leqq M_0$$

and which converges to a function $w_0(z)$ uniformly in the interior. Then $w_0(z)$ belongs to $\mathfrak{F}$ and $w^n(z)$ converges weakly to $w_0(z)$.

† Cf. Part II, §2.

1. From $\iint_{D'}|w^n|^2dxdy \leqq M_0$ we get $\iint_{D'}|w_0|^2dxdy \leqq M_0$ and therefore $w_0(z)$ belongs to $\mathfrak{F}$.

2. Let $W(z)$ be a function of $\mathfrak{F}$. Without loss of generality, we may assume $w_0(z)=0$. To a given number $\epsilon>0$ we choose a subdomain D' such that $\iint_{D-D'}|W|^2dxdy \leqq \epsilon^2$ and a number n_ϵ such that $|w^n(z)| \leqq \epsilon$ in D' for $n \geqq n_\epsilon$. Then we have for $n \geqq n_\epsilon$

$$\left|\iint_D \overline{W} w^n dxdy\right| \leqq \epsilon^2\left[\iint_D dxdy \iint_D |W|^2 dxdy + M\right]$$

and therefore $\iint_D \overline{W} w^n dxdy \to 0$ as $n \to \infty$.

Lemma 6. *A sequence $\mathfrak{S}$ of functions $w^1(z)$, $w^2(z)$, $\cdots$ in $\mathfrak{F}$ which converges weakly to zero converges to zero uniformly in the interior of D.*

The sequence $\mathfrak{S}$ is equicontinuous in the interior of D; this fact follows from (2) of Lemma 2 and $\iint_D|w^n|^2dxdy \leqq M_0$. Therefore it is sufficient to show that the sequence $\mathfrak{S}$ converges to zero at every point of D, the uniformity being a consequence of the equicontinuity.

Let z_0 be a point of D; let $\mathfrak{S}'$ be a subsequence of $\mathfrak{S}$ which at z_0 converges to a certain value w. Since $\mathfrak{S}'$ is also equicontinuous, it contains a subsequence $\mathfrak{S}''$ which converges uniformly in the interior to a certain function $w_0(z)$. According to Lemma 5, $w_0(z)$ belongs to $\mathfrak{F}$ and $\mathfrak{S}''$ converges weakly to $w_0(z)$. As $\mathfrak{S}''$ also converges weakly to zero, we have $w_0(z)=0$. Therefore $w=w_0(z_0)=0$. Thus $\mathfrak{S}$ converges to zero in every point of D and Lemma 6 is proved.

8. A characteristic value problem

In the present paragraph we discuss the form

$$w_1Vw_2 = \Re\iint_D w_1w_2dxdy$$

where $w_1(z)$ and $w_2(z)$ are functions in $\mathfrak{F}$; we observe that this form is

(i) symmetric,

$$w_1Vw_2 = w_2Vw_1;$$

(ii) bilinear,

$$wV(a_1w_1 + a_2w_2) = a_1(wVw_1) + a_2(wVw_2), \quad (a_1, a_2 \text{ real});$$

(iii) bounded,

$$|wVw| \leqq (w, w).$$

Such a form is called *completely continuous*, if, as $n\to\infty$, $w^nVw^n\to 0$ for every sequence $w^n(z)$ of functions in $\mathfrak{F}$ which converges weakly to zero.

An immediate consequence of Lemmas 4 and 5 is

THEOREM 3. *The form $\Re\iint_D w^2dxdy$ is completely continuous if the boundary has no corners except internal cusps* ($\omega=2$).

Let $w^n(z)$ be a sequence which converges weakly to zero. To prove the theorem it is sufficient to show that for every number $\epsilon>0$ we can determine a number n_ϵ such that $|\Re\iint_D(w^n)^2dxdy|\leqq 3M_0\epsilon$ for $n\geqq n_\epsilon$. To do this, we refer to Lemma 1 and the notation there used. Let S be the boundary strip of Lemma 1 corresponding to the given number ϵ, let B' be the inner boundary of S. Then, according to Lemma 6, there is a number n_ϵ such that

$$\iint_{D-S}|w^n|^2dxdy\leqq \epsilon M_0 \quad\text{and}\quad \gamma\int_{B'}|w^n|^2|dz|\leqq\epsilon M_0 \quad\text{for}\quad n\geqq n_\epsilon.$$

Now, if we note that $M=0$, in (2.1) of Lemma 1, we have

$$\left|\Re\iint(w^n)^2dxdy\right|\leqq\left|\iint_{D-S}(w^n)^2dxdy\right|+\left|\iint_S(w^n)^2dxdy\right|$$
$$\leqq\epsilon M_0+2\epsilon M_0=3\epsilon M_0.$$

Now we can apply the general theory of completely continuous forms in Hilbert spaces. This theory shows that there exist two sequences of characteristic values $\mu_1\geqq\mu_2\geqq\mu_3\geqq\cdots\to 0$ and $\mu_{-1}\leqq\mu_{-2}\leqq\cdots\to 0$ and of corresponding characteristic functions $w_1(z), w_2(z), w_3(z), \cdots, w_{-1}(z), w_{-2}(z), \cdots$ satisfying the relations

$$(w_m, w_n)=\begin{cases}1, & m=n,\\ 0, & m\neq n,\end{cases}$$

with the following properties: Let $w(z)$ be an arbitrary function of $\mathfrak{F}$. On setting

$$a_n=(w_n, w), \qquad n=\pm 1, \pm 2, \pm 3, \cdots$$

we have the relation

$$\mu_na_n=(w_nVw)$$

and the developments

$$(w, w)=a_1^2+a_2^2+\cdots+a_{-1}^2+a_{-2}^2+\cdots$$
$$(wVw)=\mu_1a_1^2+\mu_2a_2^2+\cdots+\mu_{-1}a_{-1}^2+\mu_{-2}a_{-2}^2+\cdots.$$

If for any value μ_* a function $w_*(z)=0$ in $\mathfrak{F}$ satisfies the characteristic

relation $wVw_* = \mu_*(w, w_*)$ for all w of $\mathfrak{F}$, then μ_* is contained among the values μ_n and the function $w_*(z)$ is a linear combination of the characteristic functions belonging to all $\mu_n = \mu_*$.

We observe that, simultaneously with μ_n and $w_n(z)$, $-\mu_n$ and $iw_n(z)$ also satisfy the characteristic relation. For that reason we can set

$$\mu_{-n} = -\mu_n \quad \text{and} \quad w_{-n}(z) = iw_n(z).$$

Then

$$\iint_D \overline{w_n(z)}\, w_m(z)\, dxdy = \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases}$$

and, on setting

$$c_n = a_n - ia_{-n} = \iint_D \overline{w_n(z)}\, w(z)\, dxdy,$$

we obtain

$$\iint_D |w|^2 dxdy = |c_1|^2 + |c_2|^2 + \cdots$$

$$\Re \iint_D w^2 dxdy = \Re[\mu_1 c_1^2 + \mu_2 c_2^2 + \cdots].$$

If we take $e^{i\pi/4}\, w(z)$ instead of $w(z)$, we get

$$\Im \iint_D w^2 dxdy = \Im[\mu_1 c_1^2 + \mu_2 c_2^2 + \cdots].$$

We further observe the relation

$$\iint_D |w - c_1 w_1 - c_2 w_2 - \cdots - c_n w_n|^2 dxdy$$

$$= \iint_D |w|^2 dxdy - |c_1|^2 - |c_2|^2 - \cdots - |c_n|^2 \to 0 \quad \text{and} \quad n \to \infty,$$

and we deduce from it, according to Lemma 2, that $c_1 w_1(z) + c_2 w_2(z) + \cdots$ converges to $w(z)$ uniformly in the interior. We may record these results by formulating

THEOREM 4. *If the boundary B has no corners except internal cusps, then there is a sequence of non-negative characteristic values*

$$\mu_1 \geqq \mu_2 \geqq \mu_3 \geqq \cdots \to 0$$

and a sequence of characteristic functions in $\mathfrak{F}$

$$w_1(z),\ w_2(z),\ \cdots$$

orthonormal in the sense

$$\iint_D \bar{w}_n w_m dxdy = \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases}$$

which satisfy the characteristic relation

$$\iint_D w_n w dxdy = \mu_n \iint_D \bar{w}_n w dxdy$$

for every function $w(z)$ *in* $\mathfrak{F}$. *For each function* $w(z)$ *in* $\mathfrak{F}$ *we have an expansion*

$$w(z) = c_1 w_1(z) + c_2 w_2(z) + \cdots$$

converging uniformly in the interior of D *and for the corresponding quadratic forms*

$$\iint_D |w|^2 dxdy = |c_1|^2 + |c_2|^2 + |c_3|^2 + \cdots,$$

$$\iint_D w^2 dxdy = \mu_1 c_1^2 + \mu_2 c_2^2 + \mu_3 c_3^2 + \cdots,$$

where

$$c_n = \iint_D \bar{w}_n w dxdy.$$

Since

$$\left| \iint_D w^2 dxdy \right| \leqq \iint_D |w|^2 dxdy$$

and the equality

$$\iint_D w^2 dxdy = \iint_D |w|^2 dxdy$$

actually holds if and only if w is a real constant, we may state

REMARK 1. *The first characteristic value* $\mu_1 = 1$, *and the first characteristic function* $w(z)$ *is a real constant.*

REMARK 2. *The second characteristic value is less than* 1, $\mu_2 < 1$.

This inequality and the developments of $\iint_D |w|^2 dxdy$ and $\iint_D w^2 dxdy$ confirm Theorem 2, since relation $\iint_D w dxdy = 0$ is equivalent to $c_1 = 0$ and $\mu_n \leqq \mu_2 = \theta$ for $n = 2, 3, 4, \cdots$.

9. Cases of circle and ellipse

For the circle and the ellipse the characteristic values and functions can be given explicitly.

THEOREM 5. *If D is the circle $|z|<1$, then $\mu_2=\mu_3=\cdots=0$.*

For every analytic function $W(z)$ which is regular in $|z|<1$ the mean value theorem gives

$$\iint_D W(z)dxdy = W(0).$$

Since together with $w(z)$ also $w^2(z)$ is analytic in D, we have

$$\iint_D wdxdy = w(0),$$

$$\iint_D w^2dxdy = w^2(0).$$

Consequently $c_1=0$ is equivalent to $w(0)=0$ and for all functions with $c_1=0$ we have

$$\iint_D w^2dxdy = 0.$$

THEOREM 6. *Let D be the ellipse*

$$\frac{x^2}{\cosh^2\sigma}+\frac{y^2}{\sinh^2\sigma}<1,$$

then

$$\mu_n = \frac{n\sinh 2\sigma}{\sinh 2n\sigma}, \qquad n = 1, 2, 3, \cdots$$

and the characteristic functions are the derivatives of the Tchebycheff polynomials,

$$w_n(z) = \rho_n\frac{\sinh(n \operatorname{arc}\cosh z)}{\sinh(\operatorname{arc}\cosh z)} = \rho_n\frac{2^{n-1}}{n}T_n'(z),$$

where

$$\rho_n = \left(\frac{2n}{\pi\sinh 2n\sigma}\right)^{1/2}.$$

These functions $w_1(z)$, $w_2(z)$, $w_3(z)$, $\cdots$ are polynomials of degrees 0, 1, 2, $\cdots$. The set of such polynomials is complete in the sense that every function $w(z)$ in $\mathfrak{F}$ can be approximated by a polynomial

$$p_n(z) = c_1^n w_1(z) + c_2^n w_2(z) + \cdots + c_n^n w_n(z),$$

so that

$$\iint_D |w(z) - p_n(z)|^2 dxdy$$

becomes arbitrarily small.† It remains to prove the orthogonality relation and the characteristic relation. To simplify the calculations we transform the ellipse D into a rectangular domain Δ of the $(\zeta=\xi+i\eta)$-plane. We set $z=\cosh\zeta$ and Δ: $0\leqq\xi<\sigma$, $-\pi<\eta<\pi$. Then we have

$$\iint_D \bar{w}_n w_m dxdy = \rho_n\rho_m \iint_\Delta \cosh n\zeta \cosh m\zeta d\xi d\eta = 0,$$

if $n\neq m$. Also

$$\begin{aligned}\iint_D |w_n|^2 dxdy &= \rho_n^2 \iint_\Delta |\cosh n\zeta|^2 d\xi d\eta \\ &= \rho_n^2 \int_{-\pi}^{\pi}\int_0^\sigma \frac{1}{2}\cosh 2n\xi d\xi d\eta \\ &= \rho_n^2 \pi \frac{1}{2n}\sinh 2n\sigma = 1,\end{aligned}$$

$$\begin{aligned}\iint_D w_n w_m dxdy &= \rho_n\rho_m \iint_\Delta \frac{\sinh n\zeta \sinh m\zeta}{\sinh \zeta}\overline{\sinh \zeta} d\xi d\eta \\ &= \sum\nolimits_\alpha \iint_\Delta \sinh \alpha\zeta\, \overline{\sinh \zeta} d\xi d\eta,\end{aligned}$$

where α runs through

$$\alpha = m - n + 1,\ m - n + 3, \cdots, m + n - 1, \quad \text{if} \quad m \geqq n.$$

The integral $\iint_\Delta \sinh \alpha\zeta\, \overline{\sinh \zeta}\, d\xi d\eta$ vanishes except when $\alpha = 1$, in which case its value is ρ_1^{-2}; this case occurs only, if $m=n$; therefore we have

$$\iint_D w_n w_m dxdy = \begin{cases} 0, & n \neq m, \\ \rho_n^2/\rho_1^2 = \mu_n, & n = m. \end{cases}$$

10. General case of boundary with corners

Since the form $\Re\iint_D w^2 dxdy$ is symmetric and bounded, the general theory of spectra‡ is applicable also when this form is not completely continuous,

† L. Bieberbach, *Zur Theorie und Praxis der konformen Abbildung*, Rendiconti del Circolo Matematico di Palermo, vol. 38 (1914), p. 98.

T. Carleman, *Über die Approximation analytischer Funktionen durch lineare Aggregate von vorgegebenen Potenzen*, Arkiv för Matematik, Astronomi och Fysik, vol. 17, No. 9 (1922).

For a more detailed reference, see J. L. Walsh, *Approximation by Polynomials in the Complex Domain*, Mémorial des Sciences Mathématiques, No. 73, 1935, p. 61.

‡ Cf. M. H. Stone, *Linear Transformations in Hilbert Space and their Applications to Analysis*, American Mathematical Society Colloquium Publications, vol. 15, 1932.

in which case it has a continuous spectrum. We use the notion of limit spectrum (Häufungsspektrum) in the sense of Weyl;† this closed set of values μ consists of all points of the continuous spectrum, of the limit points of the characteristic values and of all characteristic values of infinite order. In every closed interval outside of the limit spectrum there are only a finite number of characteristic values and these have a finite order. The limit spectrum consists of the single point $\mu=0$ if and only if the form is completely continuous. Now we can prove

THEOREM 7. *The l.u.b.* $\bar{\mu}$ *and the g.l.b.* $\underline{\mu}$ *of the limit spectrum of the form* $\Re\iint_D w^2dxdy$ *are precisely* M *and* $-$M *respectively. Therefore Theorem 4 is valid, if and only if* M $=0$; *that is to say, if the boundary has no corner except interior cusps.*

First we show that the limit spectrum is contained in the interval $-\mathrm{M}\leqq\mu\leqq\mathrm{M}$. From Lemma 1 we derive the inequality

$$\Re\iint_D w^2dxdy - \iint_{D-S}|w|^2dxdy - \gamma\int_{B'}|w|^2|dz| \leqq (\mathrm{M}+\epsilon)\iint_D|w|^2dxdy$$

which shows that the l.u.b. of the whole spectrum and therefore also the l.u.b. of the limit spectrum of the form of the left-hand member is not greater than $\mathrm{M}+\epsilon$. Now we refer to the fundamental theorem of Weyl‡ to the effect that the limit spectrum of a form remains unaltered whenever the form is changed by adding a completely continuous form. Now, according to Lemma 6, the forms $\iint_{D-S}|w|^2dxdy$ and $\gamma\int_{B'}|w|^2|dz|$ are completely continuous and therefore the l.u.b. of the limit spectrum of $\Re\iint_D w^2dxdy$ is not greater than M, since ϵ is arbitrarily small. The same reasoning shows that the g.l.b. is not less than $-$M. Thus we have proved that $-\mathrm{M}\leqq\underline{\mu}\leqq\bar{\mu}\leqq\mathrm{M}$.

We now prove that $\bar{\mu}\geqq\mathrm{M}$. According to a remark of H. Weyl‡ this value has the following property: Whenever there is a sequence of functions $w(z)$ in $\mathfrak{F}$ weakly convergent to zero for which

$$\Re\iint_D w^2dxdy : \iint_D|w|^2dxdy \to \mu,$$

then

$$\mu \leqq \bar{\mu}.$$

† H. Weyl, *Über beschränkte quadratische Formen, deren Differenz vollstetig ist*, Rendiconti del Circolo Matematico di Palermo, vol. 27 (1909), pp. 373–392.

‡ Loc. cit.

Let $z=0$ be a corner of the boundary B with an angle ω and with $|\sin \omega\pi/\omega\pi| = \mathrm{M}$. We set $z=re^{i\vartheta}$. We represent the two branches of B in the neighborhood of $z=0$ by $\vartheta=\vartheta_+(r)$ and $\vartheta=\vartheta_-(r)$ and choose a number $R>0$ such that the domain

$$D_R:\quad \vartheta_-(r) < \vartheta < \vartheta_+(r), \qquad 0 < r < R,$$

is contained in D.

We may assume $\vartheta_+(0)=\omega\pi/2$, $\vartheta_-(0)=-\omega\pi/2$. Then we set $\eta=1$ if $\sin \omega\pi>0$, $\eta=i$ if $\sin \omega\pi<0$ and choose the sequence

$$w_\alpha(z) = \eta(2\alpha)^{1/2}z^{\alpha-1}, \qquad \alpha \to 0.$$

We calculate $\iint_D |w_\alpha|^2 dxdy$ and $\iint_D w_\alpha^2 dxdy$. On setting

$$\vartheta_+(r) - \vartheta_-(r) = \omega\pi + r\theta(r),$$

$$\int_{\vartheta_-(r)}^{\vartheta_+(r)} e^{(2\alpha-2)i\vartheta}d\vartheta = \sin \omega\pi + r\phi(r),$$

we find that $|\theta(r)|$ and $|\phi(r)|$ are bounded for $0<r<R$. Hence, as $\alpha\to 0$, we have

$$\iint_D |w_\alpha|^2 dxdy = \omega\pi R^{2\alpha} + 2\alpha\int_0^R r^{2\alpha}\theta(r)dr + 2\alpha\iint_{D-D_R} r^{2\alpha-2}dxdy \to \omega\pi,$$

$$\iint_D w_\alpha^2\, dxdy = |\sin \omega\pi|R^{2\alpha} + 2\alpha\eta^2\int_0^R r^{2\alpha}\phi(r)dr + 2\alpha\iint_{D-D_R} z^{2\alpha-2}dxdy \to |\sin \omega\pi|,$$

and therefore

$$\iint_D w_\alpha^2\, dxdy : \iint_D |w_\alpha|^2 dxdy \to \left|\frac{\sin \omega\pi}{\omega\pi}\right| = \mathrm{M}.$$

According to Lemma 5, the sequence $w_\alpha(z)$ converges weakly to zero, as $\alpha\to 0$. Thus we obtain $\mathrm{M} \leqq \bar{\mu}$. If we take the sequence $w_\alpha(z)=i\eta\ (2\alpha)^{1/2}z^{\alpha-1}$ we get $-\mathrm{M} \geqq \mu$. So we have proved Theorem 7.

Remark. In the case where the boundary B has an external cusp which was excluded hitherto it can happen that the limit spectrum reaches the points $\mu=1$ and $\mu=-1$ and that, consequently, the inequality theorem does not hold.

We assume that the boundary B has an internal cusp at the point $z=0$. On introducing the functions $\vartheta=\vartheta_+(r)$, $\vartheta=\vartheta_-(r)$ and the domain D_R as before we assume that

$$\vartheta_+(r) = \frac{\kappa}{2} r + r^2 \theta_+(r), \qquad \vartheta_-(r) = -\frac{\kappa}{2} r + r^2 \theta_-(r),$$

where $\kappa \neq 0$ and θ_+, θ_- are bounded for $0 < r < R$. Then we choose the sequence

$$w_\alpha(z) = (2\alpha)^{1/2} z^{\alpha - 3/2}.$$

On setting

$$\vartheta_+(r) - \vartheta_-(r) = \kappa r + r^2 \theta(r),$$

$$\int_{\vartheta_-(r)}^{\vartheta_+(r)} e^{(2\alpha-3)i\vartheta} d\vartheta = \kappa r + r^2 \phi(r),$$

we find that $\theta(r)$ and $\phi(r)$ are bounded for $0 < r < R$. Hence, as $\alpha \to 0$, we have

$$\iint_D |w_\alpha|^2 dx dy = \kappa R^{2\alpha} + 2\alpha \int_0^R r^{2\alpha}\theta(r) dr + 2\alpha \iint_{D-D_R} r^{2\alpha-3} dx dy \to \kappa,$$

$$\iint_D |w|^2 dx dy = \kappa R^2 + 2\alpha \int_0^R r^{2\alpha}\phi(r) dr + 2\alpha \iint_{D-D_R} r^{2\alpha-3} dx dy \to \kappa,$$

and, therefore,

$$\iint_D w_\alpha^2 \, dx dy : \iint_D |w_\alpha|^2 dx dy \to 1.$$

Since w_α tends weakly to zero this relation yields $\bar{\mu} = 1$.

Part II. The case of two functions of two variables

This second part is concerned with complex-valued functions $k(x, y)$ of two variables x, y, defined in an open domain D. The manifold of all such functions $k(x, y)$ for which the integral

$$(k \mid k) = \frac{1}{2} \iint_D \left\{ \left| \frac{\partial k}{\partial x} \right|^2 + \left| \frac{\partial k}{\partial y} \right|^2 \right\} dx dy$$

is finite forms a Hilbert space $\mathfrak{K}$.

I consider several subspaces of $\mathfrak{K}$ and investigate their relations. By means of the projectors of these subspaces I can represent the operator which corresponds to the quadratic form treated in Part I.

The inequality and the expansion theorem which I have established for this form can be employed here. Under the same assumption regarding the boundary, I obtain an inequality and expansions for the functions k of the space $\mathfrak{K}$. This inequality statement is that there exists a positive number $\sigma = \sigma_D$ such that

$$\sigma(k \mid k) \leqq \iint_D \left(\Re \frac{\partial k}{\partial z}\right)^2 dxdy + \iint_D \left|\frac{\partial k}{\partial z}\right|^2 dxdy$$

for all functions k in $\Re$ which satisfy the relation $\Im\iint_D(\partial k/\partial z)dxdy = 0$. Finally I show that this inequality† plays a decisive part in the theory of equilibrium and vibration of an elastic plate.

1. The relative spectrum of two subspaces

1.1. The smallest angle between two spaces. At first we make some remarks on the relative spectrum of two subspaces of a Hilbert space which we will apply to our question in functional spaces.

Let $\mathfrak{H}$ be a real Hilbert space of elements h with the inner product (h_1, h_2). Let $\mathfrak{F}$ and $\mathfrak{Q}$ be two closed (linear) subspaces of $\mathfrak{H}$ with the projectors F and Q; the spaces of all elements in $\mathfrak{H}$ which are orthogonal to $\mathfrak{F}$ and $\mathfrak{Q}$ respectively are denoted by $\mathfrak{G}$ and $\mathfrak{P}$ with projectors G and P. So we have

$$F + G = 1, \qquad FG = 0, \qquad P + Q = 1, \qquad PQ = 0.$$

We denote by $\mathfrak{H}'$, $\mathfrak{F}'$, $\mathfrak{G}'$, $\mathfrak{P}'$, $\mathfrak{Q}'$ the subspaces of $\mathfrak{H}$, $\mathfrak{F}$, $\mathfrak{G}$, $\mathfrak{P}$, $\mathfrak{Q}$ which are orthogonal to the four section spaces $\mathfrak{Q}\mathfrak{F}$, $\mathfrak{P}\mathfrak{F}$, $\mathfrak{Q}\mathfrak{G}$, $\mathfrak{P}\mathfrak{G}$; so we have $\mathfrak{H}' = \mathfrak{P}' \oplus \mathfrak{Q}' = \mathfrak{F}' \oplus \mathfrak{G}'$, $\mathfrak{H} = \mathfrak{H}' \oplus \mathfrak{Q}\mathfrak{F} \oplus \mathfrak{P}\mathfrak{F} \oplus \mathfrak{Q}\mathfrak{G} \oplus \mathfrak{P}\mathfrak{G}$.

We introduce the largest non-negative number $\tau_0 \leqq \pi/2$ such that for all elements f in $\mathfrak{F}$, q in $\mathfrak{Q}$ which both are orthogonal to the section $\mathfrak{Q}\mathfrak{F}$ the relation

$$(1.1) \qquad (f, q)^2 \leqq \cos^2 \tau_0 (f, f)(q, q)$$

is valid, and we call it the "smallest angle between the spaces $\mathfrak{Q}$ and $\mathfrak{F}$." The inequality (1.1) is equivalent to each of the following four inequalities:

$$(1.2)_Q \qquad (Qf, Qf) \leqq \cos^2 \tau_0 (f, f);$$

$$(1.2)_P \qquad (Pf, Pf) \geqq \sin^2 \tau_0 (f, f);$$

$$(1.2)_F \qquad (Fq, Fq) \leqq \cos^2 \tau_0 (q, q);$$

$$(1.2)_G \qquad (Gq, Gq) \geqq \sin^2 \tau_0 (q, q).$$

We show this for $(1.2)_Q$: from (1.1) we get

$$(Qf, Qf)^2 = (f, Qf)^2 \leqq \cos^2 \tau_0 (f, f)(Qf, Qf)$$

and thus $(1.2)_Q$; and from $(1.2)_Q$ we get

† It is the analogue of the inequality of A. Korn for functions of three variables. The expansion theorem is related to those of E. and F. Cosserat.

Cf. A. Korn, *Über einige Ungleichungen, welche in der Theorie der elastischen und elektrischen Schwingungen eine Rolle spielen*, Bulletin de l'Académie des Sciences de Cracovie, 1909, vol. 2, pp. 705–724, and literature indicated therein.

$$(f, q)^2 = (Qf, q)^2 \leqq (Qf, Qf)(q, q) \leqq \cos^2 \tau_0 (f, f)(q, q)$$

and thus (1.1).

Remark. Every subspace of finite dimension has a *positive* smallest angle with respect to every other space in which it is not contained.

We have the following

THEOREM 1.1. *If the smallest angle τ_0 between the spaces $\mathfrak{Q}$ and $\mathfrak{F}$ is positive, then there is a constant ρ such that*

$$\rho(h, h) \leqq (h, Ph) + (h, Gh)$$

for all elements h in $\mathfrak{H}$ which are orthogonal to $\mathfrak{Q}\mathfrak{F}$.

For every such element h can be written in the form

$$h = f + q + j,$$

where the elements f in $\mathfrak{F}$, q in $\mathfrak{Q}$, j in $\mathfrak{P}\mathfrak{G}$ are orthogonal to $\mathfrak{Q}\mathfrak{F}$. Then we have

$$\begin{aligned}(h, h) &= (f, f) + 2(f, q) + (q, q) + (j, j)\\ &\leqq (f, f) + 2 \cos \tau_0 [(f, f)(q, q)]^{1/2} + (q, q) + (j, j)\\ &\leqq (1 + \cos \tau_0)[(f, f) + (q, q)] + (j, j)\end{aligned}$$

and

$$\begin{aligned}(h, Ph) + (h, Gh) &= (f, Pf) + (q, Gq) + 2(j, j)\\ &\geqq \sin^2 \tau_0 [(f, f) + (q, q)] + 2(j, j).\end{aligned}$$

Thus Theorem 1.1 is true with $\rho = 1 - \cos \tau_0$.

1.2. **The spectrum of the operator FQF.** Now we discuss the symmetric operator FQF. It transforms every element f in $\mathfrak{F}$ into an element $FQFf$ in $\mathfrak{F}$ and we have

$$0 \leqq (f, FQFf) \leqq (f, f).$$

The operator FQF, considered an operator in $\mathfrak{F}$, has a spectral resolution. For the sake of simplicity we assume that the spectrum is a pure point spectrum. A characteristic value κ is a real number to which there are elements $f_\kappa \neq 0$ in $\mathfrak{F}$ such that

$$FQFf_\kappa = \kappa f_\kappa.$$

The space of all such characteristic elements f_κ may be designated by $\{f_\kappa\}$; the characteristic spaces $\{f_\kappa\}$ for different values κ are orthogonal and all these spaces span the whole space $\mathfrak{F}: \sum_\kappa \{f_\kappa\} = \mathfrak{F}$. The characteristic values κ

satisfy the relation $0 \leqq \kappa \leqq 1$. The characteristic spaces of $\kappa = 1$ and $\kappa = 0$ are the sections of $\mathfrak{Q}$ and $\mathfrak{P}$ respectively with $\mathfrak{F}$:

$$\{f_1\} = \mathfrak{Q}\mathfrak{F}, \quad \{f_0\} = \mathfrak{P}\mathfrak{F}.$$

If we wish to exclude the values $\kappa = 0$ and $\kappa = 1$ we write $\sum'$ instead of $\sum$. So we have $\sum'_\kappa \{f_\kappa\} = \mathfrak{F}'$. Let f_κ be such a characteristic element. Then the element $q_\kappa = Qf_\kappa$ has the properties

$$(q_\kappa, q_\kappa) = (q_\kappa, f_\kappa) = \kappa(f_\kappa, f_\kappa).$$

Therefore, if τ is the angle between q_κ and f_κ we have

$$\kappa = \cos^2 \tau.$$

The g.l.b. of all such numbers $\tau \neq 0$ is the smallest angle τ_0 between the spaces $\mathfrak{Q}$ and $\mathfrak{F}$ and if τ_1 is the l.u.b. of all such numbers $\tau \neq \pi/2$, then $\pi/2 - \tau_1$ is the smallest angle between the spaces $\mathfrak{P}$ and $\mathfrak{F}$. The manifold which is spanned by the characteristic elements f_κ and by $q_\kappa = Qf_\kappa$, provided that $\kappa \neq 0$ and $\kappa \neq 1$, may be denoted by $\{f_\kappa, q_\kappa\}$. The dimension of $\{f_\kappa, q_\kappa\}$ is twice the multiplicity of $\{f_\kappa\}$. The subspaces $\{f_\kappa, q_\kappa\}$ for different values κ are orthogonal to each other, to $\mathfrak{P}\mathfrak{F}$, and to $\mathfrak{Q}\mathfrak{F}$. This fact follows from

$$\kappa'(f_\kappa, f_{\kappa'}) = (f_\kappa, q_{\kappa'}) = (q_\kappa, q_{\kappa'}) = (q_\kappa, f_{\kappa'}) = \kappa(f_\kappa, f_{\kappa'}).$$

These spaces are also orthogonal to the sections $\mathfrak{Q}\mathfrak{G}$ and $\mathfrak{P}\mathfrak{G}$; and we have

THEOREM 1.2. *The spaces $\{f_\kappa, q_\kappa\}$ belonging to all characteristic values $\kappa \neq 0$, $\neq 1$ of the operator FQF in $\mathfrak{F}$ span the whole space $\mathfrak{H}'$: $\mathfrak{H}' = \sum'_\kappa \{f_\kappa, q_\kappa\}$.*

Let h be an element of $\mathfrak{H}'$ which is orthogonal to all manifolds $\{f_\kappa, q_\kappa\}$, $(\kappa \neq 0, \neq 1)$. Then h is orthogonal to all $\{f_\kappa\}$. Since these elements f_κ span the whole space $\mathfrak{F}$, the element h is orthogonal to $\mathfrak{F}$; that means h belongs to $\mathfrak{G}'$. Further, the element h is orthogonal to all elements $q_\kappa = Qf_\kappa$ and, consequently, the element Qh of $\mathfrak{Q}'$ is orthogonal to $\mathfrak{F}'$; that means: Qh belongs to $\mathfrak{G}'$ and, since $\mathfrak{Q}'\mathfrak{G}' = 0$ we have $Qh = 0$. Therefore h belongs to $\mathfrak{P}'$; but, since $\mathfrak{P}'\mathfrak{G}' = 0$, we have $h = 0$. Thus Theorem 1.2 is proved.

1.3. **The spectrum of the operator $aP + bG$.** We investigate the operator

$$aP + bG,$$

where $a \neq 0$ and $b \neq 0$ are given numbers. Since this operator is symmetric and bounded it has a spectral resolution. We can express the characteristic values λ of this operator by the characteristic values $\kappa = \cos^2 \tau$ of the operator FQF in $\mathfrak{F}$. To every such value we determine the solutions λ_κ of the quadratic equation

$$(\lambda_\kappa - a)(\lambda_\kappa - b) = ab\kappa,$$

namely,

$$\lambda_\kappa^+ = \frac{a+b}{2} + \left[\left(\frac{a-b}{2}\right)^2 + ab\kappa\right]^{1/2} = \frac{a+b}{2} + \left[\left(\frac{a+b}{2}\right)^2 - ab\sin^2\tau\right]^{1/2}$$

$$= \frac{a+b}{2} - \left[\left(\frac{a-b}{2}\right)^2 + ab\kappa\right]^{1/2} = \frac{a+b}{2} - \left[\left(\frac{a+b}{2}\right)^2 - ab\sin^2\tau\right]^{1/2}.$$

If $\kappa=0$ we have $\lambda_0^+=\max(a, b)$, $\lambda_0^-=\min(a, b)$; if $\kappa=1$, $\lambda_1^+=\max(a+b, 0)$, $\lambda_1^-=\min(a+b, 0)$. We have the

THEOREM 1.3. *The values $\lambda=\lambda_\kappa^\pm$, for every $\kappa\neq 0, \neq 1$, are characteristic values of the operator $aP+bG$; their characteristic functions are*

$$h_\kappa^\pm = (\lambda_\kappa^\pm - b)f_\kappa - aq_\kappa.$$

The characteristic spaces of the values $\lambda=a$ and $\lambda=b$ are $\mathfrak{P}\mathfrak{F}$ and $\mathfrak{Q}\mathfrak{G}$ respectively and the characteristic spaces of the values $\lambda=a+b$ and $\lambda=0$ are $\mathfrak{P}\mathfrak{G}$ and $\mathfrak{Q}\mathfrak{F}$ respectively. The other characteristic functions span the whole space $\mathfrak{H}'$:

$$\mathfrak{G}' = \sum_\kappa{}'\{h_\kappa^+\} + \sum_\kappa{}'\{h_\kappa^-\}. \tag{1.3}$$

A simple calculation shows that the elements $h_\kappa^\pm$ and the four section spaces are characteristic. Since

$$\lambda_\kappa^+ - \lambda_\kappa^- \neq 0 \text{ if } \kappa \neq 0, \neq 1, \text{ we have}$$

$$f_\kappa = \frac{h^+ - h^-}{\lambda_\kappa^+ - \lambda_\kappa^-}, \qquad q_\kappa = \frac{1}{a}\,\frac{(\lambda_\kappa^+ - b)h_\kappa^+ - (\lambda_\kappa^- - b)h_\kappa^-}{\lambda_\kappa^+ - \lambda_\kappa^-}$$

and

$$\{h_\kappa^+\} \oplus \{h_\kappa^-\} = \{f_\kappa, q_\kappa\}, \qquad \kappa \neq 0 \neq 1.$$

Therefore relation (1.3) follows from Theorem 1.2 and Theorem 1.3 is proved

We remark that the spectrum is contained within the two closed intervals $[\lambda_0^-, \lambda_1^-]$ and $[\lambda_0^+, \lambda_1^+]$, which have a common point only if $a+b=0$ or $a=b$. Every limit point κ_∞ of the κ-spectrum corresponds to two limit points λ_∞^+ and λ_∞^- of the λ-spectrum (except when $a=b$, $\kappa_\infty=0$ or $a+b=0$, $\kappa_\infty=1$).

If the spaces $\mathfrak{Q}$ and $\mathfrak{F}$ have a positive smallest angle $\tau_0>0$, then there is a constant $\rho>0$ such that $\rho(h, h)\leqq a(h, Ph)+b(h, Qh)$, if $a>0$, $b>0$ for all $h\perp\mathfrak{Q}\mathfrak{F}$. (The largest possible number ρ is exactly

$$\rho_0 = \frac{a+b}{2} - \left[\left(\frac{a-b}{2}\right)^2 + ab\cos^2\tau_0\right]^{1/2}$$

$$= \frac{ab \sin^2 \tau_0}{\dfrac{a+b}{2} + \left[\left(\dfrac{a-b}{2}\right)^2 + ab \cos^2 \tau_0\right]^{1/2}}.$$

Cf. Theorem 1.1 where $a=b=1$, $\rho_0 = 1 - \cos \tau_0$.)

In the same way we can treat the spectral problem of the operator $aP+bG$ with respect to a different unit-form, for example,

$$\alpha(h_1, Fh_2) + \beta(h_1, Gh_2), \qquad \alpha \geqq 0, \qquad \beta \geqq 0.$$

In this case also every κ corresponds to two characteristic values λ, the solutions of the quadratic equation

$$(\lambda\alpha - a)(\lambda\beta - b) + (\lambda(\beta - \alpha) - b)a\kappa = 0,$$

and the characteristic elements

$$h_\kappa = (\lambda\beta - b)f_\kappa - aq_\kappa.$$

The section spaces $\mathfrak{PF}$, $\mathfrak{QG}$, $\mathfrak{PG}$, $\mathfrak{QF}$ belong to the characteristic values

$$\frac{a}{\alpha}, \quad \frac{b}{\beta}; \quad \frac{a+b}{\beta}, \quad 0.$$

2. The space $\mathfrak{K}$

2.1. **The space $\mathfrak{R}$ in general.** Let D be an open domain in the z-plane. Let $k = k_x + ik_y$ be a complex-valued function of x and y defined in D and having derivatives

$$\frac{\partial k}{\partial x}, \quad \frac{\partial k}{\partial y}$$

with respect to x and y which are L^2-integrable over any subdomain D^* of D

$$(k \mid k)_{D^*} = \frac{1}{2}\iint_{D^*}\left\{\left|\frac{\partial k}{\partial x}\right|^2 + \left|\frac{\partial k}{\partial y}\right|^2\right\} dxdy < \infty.$$

Considering k as a function of $z = x+iy$ and $\bar{z} = x - iy$ we have the relation

$$(k \mid k)_{D^*} = \iint_{D^*}\left\{\left|\frac{\partial k}{\partial z}\right|^2 + \left|\frac{\partial k}{\partial \bar{z}}\right|^2\right\} dxdy,$$

because of

$$2\frac{\partial}{\partial z} = \frac{\partial}{\partial x} - i\frac{\partial}{\partial y}, \quad 2\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}.$$

In case of a multiply-connected domain D we admit many-valued functions provided that they have singled-valued derivatives $\partial k/\partial x$, $\partial k/\partial y$.

We denote by $\mathfrak{R}$ the manifold of all functions $k(x, y)$ of this kind for which the integral

$$(k \mid k) = \iint_D \left\{ \left| \frac{\partial k}{\partial z} \right|^2 + \left| \frac{\partial k}{\partial \bar{z}} \right|^2 \right\} dxdy$$

is finite; $\mathfrak{K}$ is a linear space. Two functions k, k^* of $\mathfrak{K}$ are "equivalent" if $(k^*-k \mid k^*-k)=0$.† The manifold of all functions of $\mathfrak{K}$ which are equivalent to each other corresponds to one element of the space $\mathfrak{K}$. So, e.g., the function $k(x, y)=\text{const.}$ corresponds to the element zero.

In the space $\mathfrak{K}$ we define the inner product

$$(k_1 \mid k_2) = \mathfrak{R} \iint_D \left\{ \frac{\overline{\partial k_1}}{\partial z} \frac{\partial k_2}{\partial z} + \frac{\partial \bar{k}_1}{\partial z} \frac{\partial k_2}{\partial \bar{z}} \right\} dxdy.$$

With respect to this metric the space $\mathfrak{K}$ is a real one; but to every element $k=k_x+ik_y$ of $\mathfrak{K}$ the operations $\mathfrak{R}k=k_x$, $\mathfrak{J}k=k_y$, $ik=k_y-ik_x$, $\bar{k}=k_x-ik_y$ are feasible. Corresponding to Fischer's form of the theorem of F. Riesz and E. Fischer we have the important fact:

THEOREM 2.1. *The space* $\mathfrak{K}$ *is complete.*‡

2.2. **The subspace $\mathfrak{F}$.** We introduce the subspace $\mathfrak{F}$ of all elements f in $\mathfrak{R}$ for which $\iint_D |\partial f/\partial \bar{z}|^2 dxdy=0$; to these elements there correspond functions f with continuous derivatives for which $\partial f/\partial \bar{z}=0$, that is to say, which are analytic functions of z. In the following we assume the function f in $\mathfrak{F}$ to be analytic. Every function f of this space $\mathfrak{F}$ can also be represented by the derivative $w=df/dz$ and we have

$$(f_1 \mid f_2) = \mathfrak{R} \iint_D \bar{w}_1 w_2 dxdy. \tag{2.1}$$

Therefore the space $\mathfrak{F}$ can be identified with the space of analytic functions $w(z)$ dealt with in Part I. As we proved there, we have

† For further application we need the obvious

LEMMA 2.1. *Two functions* k, k^* *are equivalent if*

$$\iint_D |k^* - k|^2 dxdy = 0.$$

‡ For the proof use, e.g., the methods of G. Fubini, *Il principio di minimo e i teoremi di esistenza per i problemi al contorno relativi alle equazioni alle derivate parziali di ordine pari*, Rendiconti del Circolo Matematico di Palermo, vol. 23 (1907), pp. 9–11.

THEOREM 2.2. *The space $\mathfrak{F}$ is complete.*

The space $\bar{\mathfrak{F}}$ of all functions $\bar{f}$ in $\mathfrak{K}$ for which

$$\frac{\partial \bar{f}}{\partial z} = 0$$

consists of the conjugates of all functions f in $\mathfrak{F}$. We have

$$(2.2) \qquad (\bar{f}_1 \mid \bar{f}_2) = \Re \int\int_D \frac{d f_1}{d\bar{z}} \frac{d\bar{f}_2}{dz} dxdy.$$

From Theorem 2.2 we get:

THEOREM 2.3. *The space $\bar{\mathfrak{F}}$ is complete.*†

Since for every function f_1 in $\mathfrak{F}$, $\bar{f}_2$ in $\bar{\mathfrak{F}}$

$$(f_1 \mid \bar{f}_2) = 0,$$

we have

THEOREM 2.4. *The spaces $\mathfrak{F}$ and $\bar{\mathfrak{F}}$ are orthogonal to each other.*

A function $k(x, y)$ is called a potential function if it has continuous second derivatives satisfying the relation $\partial^2 k/\partial z\partial\bar{z}=0$. Such a function $k(x, y)$ can be written in the form $k(x, y)=f(z)+\overline{f^*(z)}$, where $f(z), f^*(z)$ are analytic in z; together with k also f and f^* belong to $\mathfrak{F}$. Thus we get

THEOREM 2.5. *The space $\mathfrak{F}\oplus\bar{\mathfrak{F}}$ consists of the potential functions in $\mathfrak{K}$.*

2.3. **The subspace $\mathfrak{G}$.** We introduce the subspace $\dot{\mathfrak{G}}$ of elements g in $\mathfrak{K}$ which are equivalent to functions g in $\mathfrak{K}$ which vanish identically in a boundary strip. We denote by $\mathfrak{G}$ the closure of $\dot{\mathfrak{G}}$. That is to say, $\mathfrak{G}$ consists of all functions g in $\mathfrak{K}$ for which there are functions $\dot{g}$ in $\dot{\mathfrak{G}}$ such that $(\dot{g}-g \mid \dot{g}-g)$ is arbitrarily small.‡ This definition contains the

THEOREM 2.6. *The space $\mathfrak{G}$ is complete.*

We establish the following basic identity:

$$(2.3) \qquad (k \mid g) = 2\Re \int\int_D \frac{\overline{\partial k}}{\partial z} \frac{\partial g}{\partial z} dxdy = 2\Re \int\int_D \frac{\partial \bar{k}}{\partial z} \frac{\partial g}{\partial \bar{z}} dxdy$$

for all k in $\mathfrak{K}$, g in $\mathfrak{G}$.

† This theorem is related to that of S. Zaremba. Cf. S. Zaremba, *Sur un problème toujours possible, comprenant, à titre de cas particulier, le problème de Dirichlet et celui de Neumann*, Journal de Mathématiques, sér. 9, vol. 6 (1927), pp. 127–163; O. Nikodym, *Sur un théorème de M. S. Zaremba concernant les fonctions harmoniques*, Journal de Mathématiques, sér. 9, vol. 12 (1933), pp. 95–109, and *Sur le principe du minimum*, Mathematica Cluj, vol. 9 (1936), p. 123.

‡ Under simple assumptions regarding the boundary it would be possible to give a direct definition of $\mathfrak{G}$ by a boundary condition.

To prove it let $\dot{g}$ be a function of $\dot{\mathfrak{G}}$ which vanishes in a boundary strip S. In D we take a subdomain D^* with rectifiable boundary B^* contained in S. If k is in $\mathfrak{K}$, there exists a function $k^*(x, y)$ which is defined in D^*+B^*, has continuous second derivatives, and approximates k in the sense that $(k^*-k \mid k^*-k)_{D^*}$ is small. The equations

$$\iint_{D^*} \frac{\partial \overline{k^*}}{\partial z} \frac{\partial \dot{g}}{\partial z} dxdy - \iint_{D^*} \frac{\partial \overline{k^*}}{\partial z} \frac{\partial \dot{g}}{\partial \bar{z}} dxdy$$

$$= \frac{i}{2}\int_{B^*} \frac{\partial \overline{k^*}}{\partial \bar{z}} \dot{g} d\bar{z} + \frac{i}{2}\int_{B^*} \frac{\partial \overline{k^*}}{\partial z} \dot{g} d\bar{z} = 0$$

imply that, for all functions k in $\mathfrak{K}$,

$$\iint_{D} \frac{\partial \bar{k}}{\partial z} \frac{\partial \dot{g}}{\partial z} dxdy - \iint_{D} \frac{\partial \bar{k}}{\partial z} \frac{\partial \dot{g}}{\partial \bar{z}} dxdy$$

$$= \iint_{D^*} \frac{\partial \bar{k}}{\partial z} \frac{\partial \dot{g}}{\partial z} dxdy - \iint_{D^*} \frac{\partial \bar{k}}{\partial z} \frac{\partial \dot{g}}{\partial \bar{z}} dxdy = 0.$$

Since $\dot{\mathfrak{G}}$ is dense in $\mathfrak{G}$ we conclude that (2.3) also holds.

From this identity we immediately get

THEOREM 2.7. *The space* $\mathfrak{G}$ *is orthogonal to* $\mathfrak{F}$ *and to* $\overline{\mathfrak{F}}$: $\mathfrak{G} \perp \mathfrak{F}$, $\mathfrak{G} \perp \overline{\mathfrak{F}}$.

Further we prove the decisive

THEOREM 2.8. *The spaces* $\mathfrak{F}$, $\overline{\mathfrak{F}}$, $\mathfrak{G}$ *span the whole space* $\mathfrak{K}$: $\mathfrak{F} \oplus \overline{\mathfrak{F}} \oplus \mathfrak{G} = \mathfrak{H}$.

We construct† the following function $\dot{g}$ in $\dot{\mathfrak{G}}$:

Let $|z-z_0| \leqq R$ be a circle within D, z' a point within this circle and $|z-z'| \leqq r$ a circle in the interior of $|z-z_0| \leqq R$. Then we put

$$\dot{g}(x, y) = \begin{cases} 0 & \text{for } |z - z_0| \geqq R; \\ \dfrac{1}{\pi} \log \left| \dfrac{R(z - z')}{R^2 - \overline{(z' - z_0)}(z - z_0)} \right| & \text{for } |z - z_0| \leqq R, \\ & \quad \text{but } |z - z'| \geqq r; \\ \dfrac{1}{\pi} \log \left| \dfrac{Rr}{R^2 - \overline{(z' - z_0)}(z - z_0)} \right| & \text{for } |z - z'| \leqq r. \end{cases}$$

Let k be a function orthogonal to $\mathfrak{G}$. Then from $(ig \mid k)=0$ and $(g \mid k)=0$ we

† For the following reasoning cf. G. Fubini, loc. cit., p. 10, §§6, 7, and R. Courant, *Über direkte Methoden, bei Variations- und Randwertproblemen*, Jahresbericht der Deutschen mathematiker vereinigung, vol. 34 (1925), pp. 107, 108.

have $\iint_D(dk/dz)(d\bar{g}/d\bar{z})dxdy=0$, according to (2.3). Hence we obtain by integration by parts

$$\frac{1}{2\pi i}\int_{|z-z'|=r} k(x, y)\frac{dz}{z-z'} = \frac{1}{2\pi i}\int_{|z-z_0|=R} k(x, y)\frac{R^2-|z'-z_0|^2}{|z-z'|^2}\frac{dz}{z-z_0}.$$

Consequently the mean value at the left-hand side is independent of r and is a potential function $k(x', y')$ in x', y'. The function $k^*(x, y)-k(x, y)$ has the property that

$$\frac{1}{2\pi i}\int_{|z-z'|=r}(k^*-k)\frac{dz}{z-z'} = 0$$

and, consequently,

$$\iint_{|z-z'|\leqq r}(k^*-k)dxdy = 0.$$

Therefore† $\iint_D|k^*-k|^2dxdy=0$ and k is equivalent to the potential function k^* (cf. Lemma 2.1). According to Theorem 2.5 the element k belongs to $\mathfrak{F}\oplus\bar{\mathfrak{F}}$. Thus $\mathfrak{K}=\mathfrak{F}\oplus\bar{\mathfrak{F}}\oplus\mathfrak{G}$ as we wished to prove.

We denote by F, $\bar{F}$, G the orthogonal projectors which belong to the closed subspaces $\mathfrak{F}$, $\bar{\mathfrak{F}}$, $\mathfrak{G}$; the projections of a function k of $\mathfrak{K}$ on these spaces are Fk, $\bar{F}k$, Gk respectively. Theorem 2.8 gives the relation

$$F+\bar{F}+G = 1.$$

3. The space $\mathfrak{H}$

3.1. The metric (,). In this section we deal with the space

$$\mathfrak{H} = \mathfrak{F}+\mathfrak{G}$$

consisting of all elements h of $\mathfrak{K}$ which are orthogonal to $\bar{\mathfrak{F}}$. We employ the abbreviation

$$(k_1, k_2) = \Re\iint_D \overline{\frac{\partial k_1}{\partial z}}\,\frac{\partial k_2}{\partial z}\,dxdy$$

$$(k_1; k_2) = \Re\iint_D \frac{\partial \bar{k}_1}{\partial z}\,\frac{\partial k_2}{\partial \bar{z}}\,dxdy$$

for all k in $\mathfrak{K}$, so that

† We make use of the

Lemma 2.2. *Whenever an L^2-integrable function $\phi(x, y)$ has the property that*

$$\frac{1}{r^2\pi}\iint_{|z'-z|\leqq r}\phi(x', y')dx'dy' = 0$$

for every circle $|z-z'|\leqq r$ within D, then $\iint_D|\phi|^2dxdy=0$.

$$(k_1, k_2) + (k_1; k_2) = (k_1 \mid k_2).$$

The inner product (h_1, h_2) defines a new metric $(,)$ in the space $\mathfrak{H}$, since the form (h, h) is positive definite and vanishes for h in $\mathfrak{H}$ only if $h = 0$. An immediate consequence of identities (2.1) and (2.3) is

THEOREM 3.1. *The spaces $\mathfrak{F}$ and $\mathfrak{G}$ are orthogonal with respect to the metric* $(,)$.

Further we use

LEMMA 3.1. *For elements h in $\mathfrak{H}$ the inequality*

$$\frac{1}{2}(h \mid h) \leqq (h, h) \leqq (h \mid h)$$

is valid.

From the identities (2.2) and (2.3), on writing $h = Fh + Gh = f + g$, we get in fact the relation

$$\frac{1}{2}(h \mid h) = \frac{1}{2}(f \mid f) + \frac{1}{2}(g \mid g) = \frac{1}{2}(f, f) + (g, g)$$
$$\leqq (f, f) + (g, g) = (h, h) \leqq (h, h) + (h; h) = (h \mid h).$$

From this inequality we deduce

LEMMA 3.2. *A subspace of $\mathfrak{H}$ is complete with respect to the metric $(\mid)$ if and only if it is complete with respect to the metric $(,)$.*

Thus we obtain

THEOREM 3.2. *The spaces $\mathfrak{F}$ and $\mathfrak{G}$ are closed with respect to* $(,)$.

In the following part of §3 the terms "orthogonal" and "closed" refer to the metric $(,)$ only.

The projectors F and G for the spaces $\mathfrak{F}$ and $\mathfrak{G}$ satisfy the relation

$$F + G = 1 \quad \text{in} \quad \mathfrak{H}.$$

They belong to the form $(,) - (;)$ and $(;)$ in the sense that

$$(h^*, Fh) = (h^*, h) - (h^*; h), \qquad (h^*, Gh) = (h^*; h)$$

for h^*, h in $\mathfrak{H}$, in accordance with (2.3).

3.2. **The spaces $\mathfrak{P}$ and $\mathfrak{Q}$.** We introduce the symmetric forms

$$h_1 P h_2 = \frac{1}{4} \iint_D \left(\overline{\frac{\partial h_1}{\partial z}} + \frac{\partial h_1}{\partial z} \right) \left(\frac{\partial h_2}{\partial z} + \overline{\frac{\partial h_2}{\partial z}} \right) dxdy = \iint_D \Re \frac{\partial h_1}{\partial z} \Re \frac{\partial h_2}{\partial z} \, dxdy$$

$$h_1\boldsymbol{Q}h_2 = \frac{1}{4}\iint_D\left(\overline{\frac{\partial h_1}{\partial z}} - \frac{\partial h_1}{\partial z}\right)\left(\frac{\partial h_2}{\partial z} - \overline{\frac{\partial h_2}{\partial z}}\right)dxdy = \iint_D \Im\frac{\partial h_1}{\partial z}\,\Im\frac{\partial h_2}{\partial z}\,dxdy$$

defined for h_1, h_2 in $\mathfrak{H}$. We have

$$0 \leqq h\boldsymbol{P}h \leqq (h, h), \qquad 0 \leqq h\boldsymbol{Q}h \leqq (h, h)$$

and

$$\boldsymbol{P} + \boldsymbol{Q} = (,).$$

The forms $\boldsymbol{P}$ and $\boldsymbol{Q}$ correspond to bounded symmetric operators P, Q such that

$$h^*\boldsymbol{P}h = (h^*, Ph); \qquad h^*\boldsymbol{Q}h = (h^*, Qh)$$

for h^*, h in $\mathfrak{H}$. P and Q satisfy the relation

$$P + Q = 1.$$

By $\mathfrak{P}$ and $\mathfrak{Q}$ we denote the subspaces of all the elements p and q in $\mathfrak{H}$ for which

$$Qp = 0 \quad \text{and} \quad Pq = 0,$$

respectively.

The functions

$$p = p_x + ip_y \text{ in } \mathfrak{P} \quad \text{and} \quad q = q_x + iq_y \text{ in } \mathfrak{Q}$$

can be characterized as well by

$$p\boldsymbol{Q}p = \iint_D\left(\Im\frac{\partial p}{\partial z}\right)^2 dxdy = \frac{1}{4}\iint_D\left(\frac{\partial p_y}{\partial x} - \frac{\partial p_x}{\partial y}\right)^2 dxdy = 0$$

and

$$q\boldsymbol{P}q = \iint_D\left(\Re\frac{\partial q}{\partial z}\right)^2 dxdy = \frac{1}{4}\iint_D\left(\frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y}\right)^2 dxdy = 0$$

respectively. We have

THEOREM 3.3. *The spaces* $\mathfrak{P}$ *and* $\mathfrak{Q}$ *are closed.*

Since p in $\mathfrak{P}$, q in $\mathfrak{Q}$

$$(p, q) = p\boldsymbol{P}q + p\boldsymbol{Q}q = (p, Pq) + (Qp, q) = 0$$

we have

THEOREM 3.4. *The spaces* $\mathfrak{P}$ *and* $\mathfrak{Q}$ *are orthogonal*: $\mathfrak{P} \perp \mathfrak{Q}$.

We now prove the basic

THEOREM 3.5. *The spaces* $\mathfrak{P}$ *and* $\mathfrak{Q}$ *span the whole space* $\mathfrak{H}$: $\mathfrak{P} \oplus \mathfrak{Q} = \mathfrak{H}$.

Let h be a function in $\mathfrak{H}$ which is orthogonal to $\mathfrak{P}$. Then we have

$$pPh = pPh + pQh = (p, h) = 0. \tag{3.1}$$

We take the function

$$k(x, y) = \begin{cases} z - z_0 & \text{in} \quad |z - z_0| \leqq R, \\ \dfrac{R^2}{\overline{z - z_0}} & \text{in} \quad |z - z_0| \geqq R, \end{cases}$$

which belongs to $\mathfrak{K}$ and has the property

$$\frac{\partial k}{\partial z} = 1 \quad \text{in} \quad |z - z_0| < R, \qquad = 0 \quad \text{in} \quad |z - z_0| > R.$$

The projection of this function on the space $\mathfrak{H}$: $p = k - \overline{F}k$ also has the property

$$\frac{\partial p}{\partial z} = 1 \quad \text{in} \quad |z - z_0| < R, \qquad = 0 \quad \text{in} \quad |z - z_0| > R$$

and, therefore, belongs to $\mathfrak{P}$. By inserting this function p into the relation (3.1) we get

$$\Re \iint_{|z-z_0|<R} \frac{\partial h}{\partial z} dxdy = 0.$$

Since the circle $|z-z_0| < R$ was arbitrary within D, we deduce (cf. Lemma 2.2)

$$\iint_D \left(\Re \frac{\partial h}{\partial z}\right)^2 dxdy = 0.$$

Hence h belongs to $\mathfrak{Q}$ and consequently $\mathfrak{P} \oplus \mathfrak{Q} = \mathfrak{H}$.

From Theorem 3.5 we see that the elements Qh belong to $\mathfrak{Q}$ because they are orthogonal to $\mathfrak{P}$: $(p, Qh) = (Qp, h) = 0$. Thus we have $PQ = 0$ or $Q^2 = Q$; in the same way we find $P^2 = P$. Therefore we have

THEOREM 3.6. *The operators P and Q are the projectors of the spaces $\mathfrak{P}$ and $\mathfrak{Q}$ respectively.*

Hence we see that the theory of §1 is applicable.

3.3. **The section spaces.** Before going into detail we investigate the section spaces $\mathfrak{P}\mathfrak{F}$, $\mathfrak{Q}\mathfrak{F}$, $\mathfrak{P}\mathfrak{G}$, $\mathfrak{Q}\mathfrak{G}$.

We have

$$\mathfrak{P}\mathfrak{F} = \{z\}, \qquad \mathfrak{Q}\mathfrak{F} = \{iz\},$$

for $f=z+\text{const.}$ and $f=iz+\text{const.}$ are the only functions of $\mathfrak{F}$ for which df/dz is real and imaginary respectively. The spaces $\mathfrak{P}\mathfrak{G}$ and $\mathfrak{Q}\mathfrak{G}$ consist of all functions of $\mathfrak{P}$ and $\mathfrak{Q}$, respectively, which are constant at the boundary in the approximate sense of the definition of $\mathfrak{G}$ (cf. §2). Let $\mathfrak{F}'$, $\mathfrak{G}'$, $\mathfrak{P}'$, $\mathfrak{Q}'$ be the subspaces of all functions in $\mathfrak{F}$, $\mathfrak{G}$, $\mathfrak{P}$, $\mathfrak{Q}$ which are orthogonal to the section spaces; then we have $\mathfrak{H}'=\mathfrak{F}'\oplus\mathfrak{G}'=\mathfrak{P}'\oplus\mathfrak{Q}'$. We establish the following theorem, but we shall not make use of it.

THEOREM 3.7. *The elements h of $\mathfrak{H}'$ are equivalent to functions of the form*

$$h = z\overline{\phi_1(z)} + \phi_2(z) + \overline{\phi_0(z)},$$

where $\phi_1(z)$, $\phi_2(z)$, and $\phi_0(z)$ are analytic in z.

To prove this we take a circle $|z-z_0|\leqq 2R$ in D; we choose a real function $\phi(x, y)$ which is four times continuously differentiable and which is $=1$ in $|z-z_0|\leqq R$, $=0$ in $|z-z_0|\geqq 2R$. Then we choose a number $r<R$ and a point z' of $|z'-z_0|<R-r$ and set

$$h^0 = \begin{cases} \dfrac{z-z'}{r^2} & \text{in} \quad |z-z'| \leqq r, \\[2ex] \dfrac{\partial}{\partial \bar z}\phi \log |z-z'|^2 & \text{in} \quad |z-z'| \geqq r. \end{cases}$$

This function h^0 belongs to $\mathfrak{P}\mathfrak{G}$ and ih^0 belongs to $\mathfrak{Q}\mathfrak{G}$. Therefore $(h^0,h)=(ih^0, h)=0$. This implies

$$\frac{1}{r^2\pi}\iint_{|z-z'|\leqq r}\frac{\partial h}{\partial z}dxdy = -\frac{1}{\pi}\iint_{R\leqq|z-z_0|\leqq 2R}\frac{\partial^2}{\partial z\partial\bar z}\phi\log|z-z'|^2\frac{\partial h}{\partial z}dxdy.$$

The integral on the right is a potential function in x', y' which is independent of r and can be written in the form $\overline{\phi_1(z')}+\phi_2'(z')$, where $\phi_1(z)$ and $\phi_2(z)$ are analytic in z. On setting

$$\phi^* = z\overline{\phi_1(z)} + \phi_2(z)$$

we get

$$\iint_{|z-z'|\leqq r}\frac{\partial}{\partial z}(h-\phi^*)dxdy = 0$$

and (cf. Lemma 2.2)

$$\iint_D \left|\frac{\partial(h-\phi^*)}{\partial z}\right|^2 dxdy = 0.$$

Therefore $\overline{h-\phi^*}$ is equivalent to an analytic function $\phi_0(z)$, and h itself is equivalent to $z\,\overline{\phi_1(z)}+\phi_2(z)+\overline{\phi_0(z)}$.

3.4. **The operator** M. In §1 we investigated the operator FQF in $\mathfrak{F}$. In its place, we now consider the bounded symmetric operator

$$M = F(P - Q)F = F - 2FQF,$$

which takes every element f of $\mathfrak{F}$ into the element Mf in $\mathfrak{F}$. This operator M in $\mathfrak{F}$ belongs to the form

$$f_1\boldsymbol{M}f_2 = f_1\boldsymbol{P}f_2 - f_1\boldsymbol{Q}f_2 = \Re\iint_D \frac{df_1}{dz}\frac{df_2}{dz}\,dxdy$$

in the sense that

$$f_1\boldsymbol{M}f_2 = (f_1, Mf_2).$$

Now, this form $\boldsymbol{M}$ is exactly the form dealt with in Part I if $w=df/dz$; and thus we obtain a very simple representation of the operator of this form. On the other hand we can use the properties of this form established in Part I to investigate the new forms $\boldsymbol{P}$ and $\boldsymbol{Q}$.

4. The form $\boldsymbol{E}$

4.1. **The inequality.** In this section we assume, as in Part I, that the boundary B of the domain D consists of a finite number of curves $z=z(s)$ with continuous tangent $z(s)$ except at a finite number of corners where $\dot{z}(s)$ is continuous on each side and arc $\dot{z}(s)$ jumps by less than π. Then we get (in the sense of §1)

THEOREM 4.1. *The spaces* $\mathfrak{F}$ *and* $\mathfrak{Q}$ *(or* $\mathfrak{P}$*) have a positive smallest angle.*

Under our assumption on the boundary we can deduce from Theorem 1, Part I that there is a constant $\theta<1$ such that for all functions f in $\mathfrak{F}$ satisfying the relation

$$\Re\iint_D \frac{df}{dz}\,dxdy = 0$$

the inequality

$$f\boldsymbol{M}f = \Re\iint_D\left(\frac{df}{dz}\right)^2 dxdy \leqq \theta\iint_D\left|\frac{df}{dz}\right|^2 dxdy$$

is valid. Since the form $\boldsymbol{M}$ at the left-hand side belongs to the operator $M=1-2FQF$ in $\mathfrak{F}$ we get, on setting if instead of f, the inequality

$$-(f, f) + 2(f, Qf) \leqq \theta(f, f)$$

or

$$(f, Qf) \leqq \frac{1+\theta}{2}(f, f),$$

for all functions f in $\mathfrak{F}$ which satisfy the relation

$$(iz, f) = 0$$

and which are therefore orthogonal to $\mathfrak{Q}\mathfrak{F} = \{iz\}$. (In the same way we get $(f, Pf) \leqq (1+\theta)/2\ (f, f)$ under $(z, f) = 0$.) Therefore $\cos^2 \tau_0 \leqq (1+\theta)/2 < 1$ and the smallest angle $\tau_0 < \pi/2$ is positive as we stated in Theorem 4.1.

We introduce the form

$$\begin{aligned} h_1 \boldsymbol{E} h_2 &= a(h_1, Ph_2) + b(h_1, Gh_2) = a(h_1 \boldsymbol{P} h_2) + b(h_1; h_2) \\ &\equiv a \iint_D \Re \frac{\partial h_1}{\partial z} \Re \frac{\partial h_2}{\partial z} dxdy + b\Re \iint_D \frac{\partial \bar{h}_1}{\partial z} \frac{\partial h_2}{\partial \bar{z}} dxdy \end{aligned}$$

for h in $\mathfrak{H}$. We assume a and b to be positive. Applying Theorem 1.1 of §1 we deduce immediately from Theorem 4.1 the

THEOREM 4.2. *There is a positive constant* ρ *such that*

$$\rho(h, h) \leqq (h\boldsymbol{E}h)$$

or

$$\rho \iint_D \left| \frac{\partial h}{\partial z} \right|^2 dxdy \leqq a \iint_D \left| \Re \frac{\partial h}{\partial z} \right|^2 dxdy + b \iint_D \left| \frac{\partial h}{\partial \bar{z}} \right|^2 dxdy$$

for all functions h *in* $\mathfrak{K}$ *which satisfy the condition*

$$(iz, h) = \Im \iint_D \frac{\partial h}{\partial z} dxdy = 0.$$

We now consider the form

$$k_1 \boldsymbol{E} k_2 = a \iint_D \Re \frac{\partial k_1}{\partial z} \Re \frac{\partial k_2}{\partial z} dxdy + b\Re \iint_D \frac{\partial \bar{k}_1}{\partial z} \frac{\partial k_2}{\partial \bar{z}} dxdy$$

for k in $\mathfrak{R}$; $a > 0$, $b > 0$. On setting $k = (F+G)k + \overline{F}k = h + \bar{f}$, we get

$$\begin{aligned} k_1 \boldsymbol{E} k_2 &= h_1 \boldsymbol{E} h_2 + b(\bar{f}_1 \mid \bar{f}_2) \\ (k_1 \mid k_2) &= (h_1 \mid h_2) + (\bar{f}_1 \mid \bar{f}_2) \\ &\leqq 2(h_1, h_2) + (\bar{f}_1 \mid \bar{f}_2) \end{aligned}$$

(cf. Lemma 3.1), and

$$(iz \mid k) = (iz, h).$$

Therefore we obtain from Theorem 4.2

THEOREM 4.3. *There is a positive constant σ such that*

$$\sigma(k \mid k) \leqq (k\boldsymbol{E}k),$$

or

$$\sigma \iint_D \left|\frac{\partial k}{\partial z}\right|^2 dxdy + \sigma \iint_D \left|\frac{\partial k}{\partial \bar{z}}\right|^2 dxdy \leqq a \iint_D \left(\Re \frac{\partial k}{\partial z}\right)^2 + b \iint_D \left|\frac{\partial k}{\partial \bar{z}}\right|^2 dxdy,$$

for all functions k in $\mathfrak{R}$ which satisfy the condition

$$(iz \mid k) = \Im \iint_D \frac{\partial k}{\partial z} dxdy = 0.$$

4.2. **The spectral resolution.** It is very simple to give the spectral resolution of the form $\boldsymbol{E}$ with respect to the unit-form (,) or (|) if we assume that the boundary B has no corners.

First we remark that, according to Theorem 4 of Part I, the operator $M = 1 - 2FQF$ in $\mathfrak{F}$ has a pure point spectrum of values

$$\mu_1 = 1 > \mu_2 \geqq \mu_3 \geqq \cdots \to 0; \qquad \mu_{-n} = -\mu_n.$$

Therefore the operator FQF in $\mathfrak{F}$ also has a pure point spectrum of values

$$\kappa_n = \frac{1 - \mu_n}{2} \to \frac{1}{2}, \qquad n = \pm 1, \pm 2, \pm 3, \cdots.$$

The characteristic functions of FQF and M are the same $f_n(z)$; $f_{-n}(z) = if_n(z)$. $f_1(z) = az$, $f_{-1}(z) = iaz$; a = real constant.

We now observe that the form $\boldsymbol{E}$ in $\mathfrak{H}$ belongs to the operator $aP + bG$ with respect to the unit-form (,). Therefore we can apply Theorem 1.3 of §1. We find that this operator has a pure point spectrum of values λ_n^+, λ_n^-, $n = \pm 1, \pm 2, \cdots$, which are the solutions of the quadratic equation

$$(\lambda_n - a)(\lambda_n - b) = ab \frac{1 - \mu_n}{2}.$$

For $n = \pm 1$ we get the characteristic values $\lambda = 0, a, b, a + b$; their characteristic spaces are

$$\mathfrak{Q}\mathfrak{F} = \{iz\}, \qquad \mathfrak{P}\mathfrak{F} = \{z\}, \qquad \mathfrak{Q}\mathfrak{G}, \qquad \mathfrak{P}\mathfrak{G}.$$

The characteristic functions of λ_n for $n = \pm 2, \pm 3, \cdots$ are

$$h_n = (\lambda_n - b)f_n - aQf_n.$$

They belong to the space $\mathfrak{H}'$ and span it.

The set of the characteristic values λ_n^+, λ_n^- has the two limit points

$$\lambda_\infty^+ = \frac{a+b}{2} + \frac{1}{2}\,[a^2+b^2]^{1/2}, \qquad \lambda_\infty^- = \frac{a+b}{2} - \frac{1}{2}\,[a^2+b^2]^{1/2}.$$

To give the spectral resolution of the form $k_1\boldsymbol{E}k_2$ with respect to the unit-form $(\,|\,)$ we observe the relations

$$k_1\boldsymbol{E}k_2 = h_1\boldsymbol{E}h_2 + b(\bar{f}_1\,|\,\bar{f}_2),$$
$$(k_1\,|\,k_2) = (h_1, Fh_2) + 2(h_1, Gh_2) + (\bar{f}_1\,|\,\bar{f}_2).$$

Therefore the space $\overline{\mathfrak{F}}$ is characteristic with value b. The other characteristic functions belong to $\mathfrak{H}$; to find them we apply the remarks at the end of §1 and get the following:

The values $\lambda=0$, a, $b/2$, $(a+b)/2$, b are characteristic with spaces $\mathfrak{Q}\mathfrak{F}=\{iz\}$, $\mathfrak{P}\mathfrak{F}=\{z\}$, $\mathfrak{Q}\mathfrak{G}$, $\mathfrak{P}\mathfrak{G}$, $\mathfrak{F}$. The other characteristic values λ_n^+, λ_n^-, $n=\pm 2, \pm 3, \cdots$, are the solutions of the equation

$$(\lambda_n - a)(2\lambda_n - b) + (\lambda_n - b)a\,\frac{1-\mu_n}{2};$$

they have two limit points λ_∞^+, λ_∞^-, solutions of

$$(\lambda_\infty - a)(2\lambda_\infty - b) + (\lambda_\infty - b)\,\frac{a}{2} = 0.$$

Their characteristic functions belong to the space $\mathfrak{H}'$ and span it.

5. Application to the theory of the elastic plate

Finally we outline the application of the inequality of Theorem 4.2 to the theory of an elastic plate. Let us imagine an elastic plate spread out over the domain D. We assume that the boundary B of D consists of a finite number of curves with a continuous tangent except at a finite number of corners as in §4. Let $k(x, y)$ be the displacement transforming every point z of D into the point $z+k$. Then the potential energy arising from this deformation† is

† The theory of *transversal* displacements also depends on the form $k\boldsymbol{E}k$ provided that k is the gradient $k=\partial j/\partial z$ of a real function $j(x, y)$. The equilibrium problem has been treated several times with the help of variational methods. For the first time by G. Fubini, loc. cit., and by W. Ritz, Journal für die reine und angewandte Mathematik, vol. 135 (1909). For the vibrations see W. Ritz, Annalen der Physik (1909). For both see K. Friedrichs, Mathematische Annalen, vol. 98 (1927), pp. 206–247. For these problems, where $k=\partial j/\partial z$ belongs to $\mathfrak{P}+\overline{\mathfrak{F}}$, our inequality (Theorem 4.2) need not be used if $b\neq 0$.

$$kEk = a\iint_D \left(\Re\frac{\partial k}{\partial z}\right)^2 dxdy + b\iint_D \left|\frac{\partial k}{\partial z}\right|^2 dxdy,$$

where a and b are positive constants.†

We restrict k to the subspace $\mathfrak{K}_0$ of one-valued functions in $\mathfrak{K}$. In $\mathfrak{K}_0$ we introduce the form

$$k_1Hk_2 = \Re\iint_D \bar{k}_1k_2dxdy.$$

Then the kinetic energy arising from the vibration $ke^{i\omega t}$ is proportional to

$$kHk = \iint_D |k|^2dxdy.$$

Let the complex-valued function $\phi(x, y)$ be the density of a force applied over the interior of D; the potential energy of this force is

$$-kH\phi = -\Re\iint_D \bar{k}\phi dxdy.$$

There are different cases according as the displacement k has to satisfy boundary conditions or not.

First we take the case (1) of no displacement at the boundary B. Then the displacement g belongs to the subspace $\mathfrak{G}$.

The problem of equilibrium (1) is: to find a displacement $\underline{g}$ in $\mathfrak{G}$ such that

$$2g'E\underline{g} - g'H\phi = 0$$

for all g' in $\mathfrak{G}$.

The theory of vibrations (1) requires the simultaneous spectral resolution of the forms gEg and gHg.

Now, for g in $\mathfrak{G}$ we have the relation

$$gEg \geqq b\iint_D \left|\frac{\partial g}{\partial \bar{z}}\right|^2 dxdy = \frac{b}{2}(g \mid g).$$

Since the form H in $\mathfrak{G}$ is bounded and completely continuous with respect to $(\,|\,)$, it has like properties with respect to E. Therefore no difficulty arises in solving the two problems.

In case (2) there is no boundary condition for the displacement (and no force working at the boundary).

† It is

$$a = \frac{m+1}{m-1}G = \frac{2m}{m-1}E, \qquad b = G = \frac{2m}{m-1}E$$

where E, G, m are the moduls of elasticity.

The question of *equilibrium* (2) is: to find a displacement $\underline{k}$ in $\mathfrak{K}_0$ such that

$$2k'\boldsymbol{E}\underline{k} - k'\boldsymbol{H}\phi = 0$$

for all k' in $\mathfrak{K}'$. Here we must assume $1\boldsymbol{H}\phi=0$, $i\boldsymbol{H}\phi=0$, that is, $\iint_D \phi dx=0$ and $iz\boldsymbol{H}\phi=\mathfrak{I}\iint_D \bar{z}\phi dxdy=0$ as we shall see.

The theory of *vibrations* (2) requires the simultaneous spectral resolution of the forms $k\boldsymbol{E}k$ and $k\boldsymbol{H}k$ for k in $\mathfrak{K}_0$.

In the case (2) the energy $k\boldsymbol{E}k$ vanishes for a pure translation $k=$const. and a pure rotation $k=iz$. To exclude this we take the accessory conditions

$$(5.1) \qquad 1\boldsymbol{H}k = 0, \qquad i\boldsymbol{H}k = 0 \qquad \text{or} \qquad \iint_D k\,dxdy = 0$$

$$(5.2) \qquad iz\boldsymbol{H}k = 0 \qquad \text{or} \qquad \mathfrak{I}\iint_D \bar{z}k\,dxdy = 0.$$

By $\mathfrak{K}_1$ or $\mathfrak{K}_2$ we denote the space of all functions k in $\mathfrak{K}_0$ satisfying (5.1) or (5.1) and (5.2) respectively. In $\mathfrak{K}_1$ and $\mathfrak{K}_2$ we use the unit-form $(|)$.

We employ the inequality of Poincaré:†

There is a positive constant π such that for all k in $\mathfrak{R}_1$

$$k\boldsymbol{H}k \leqq \pi(k \mid k).$$

So the form $\boldsymbol{H}$ is bounded in $\mathfrak{R}_1$. Therefore the spaces $\mathfrak{K}_1$ and $\mathfrak{K}_2$ are closed with respect to $(|)$.

Further we note†

The form $\boldsymbol{H}$ is completely continuous in $\mathfrak{K}_1$, with respect to $(|)$.

Now we prove

THEOREM 5.1. *There is a positive constant ϵ such that for all functions k in $\mathfrak{K}_2$*

$$\epsilon(k \mid k) \leqq (k\boldsymbol{E}k).$$

Since the space $\{iz\}$ has the dimension 1 it has a positive smallest angle $\tau_0 \leqq \pi/2$ with respect to the space $\mathfrak{K}_2$ and the section $\{iz\}\mathfrak{K}_2=0$. Let aiz be the projection of k into $\{iz\}$; then we have (cf. $(1.2)_P$) for k in $\mathfrak{K}_2$

$$(k - aiz \mid k - aiz) \geqq \sin^2 \tau_0(k \mid k).$$

Now, $k-aiz$ being orthogonal to $\{iz\}$, Theorem 4.3 gives a constant $\sigma>0$ such that

† Cf. e.g. K. Friedrichs, *Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die spektral Zerlegung von Differentialoperatoren*. II. Mathematische Annalen, vol. 109 (1934), pp. 705–707.

$$\sigma(k - aiz \mid k - aiz) \leqq (k - aiz)\boldsymbol{E}(k - aiz).$$

But we have

$$(k - aiz)\boldsymbol{E}(k - aiz) = (k\boldsymbol{E}k).$$

Thus Theorem 5.1 is proved.

In consequence of Theorem 5.1 we can take $\boldsymbol{E}$ as unit-form in the space $\mathfrak{K}_2$, and we are sure that every subspace of $\mathfrak{K}_2$ which is complete with respect to $(|)$ is complete with respect to $\boldsymbol{E}$; especially $\mathfrak{K}_2$ itself is closed as to $(|)$; further that every form in $\mathfrak{K}_2$ which is bounded or completely continuous with respect to $(|)$ has a like property with respect to $\boldsymbol{E}$.

Now, since $k\boldsymbol{H}\phi$ is a bounded linear form, it is a well known fact that there is a function $\underline{k}$ in $\mathfrak{R}_2$ such that

$$2k'\boldsymbol{E}\underline{k} - k'\boldsymbol{H}\phi = 0$$

for all k' in $\mathfrak{R}_2$; but since we have assumed $(a+ib)\boldsymbol{H}\phi=0$, $iz\boldsymbol{H}\phi=0$, this relation holds for all k in $\mathfrak{R}_0$. This function $\underline{k}$ is the solution of the *equilibrium problem* (2).

The form $\boldsymbol{H}$ is completely continuous with respect to $\boldsymbol{E}$. From this we get the solution of the *vibration problem* (2):

There is a sequence of values $\eta_1 \leqq \eta_2 \leqq \eta_3 \leqq \cdots \to \infty$ and of functions $k_1, k_2, k_3, \cdots$ in $\mathfrak{R}_2$ orthogonal with respect to $\boldsymbol{H}$ and $\boldsymbol{E}$ such that every function k in $\mathfrak{R}_2$ can be developed into a series

$$k = a_1k_1 + a_2k_2 + a_3k_3 + \cdots$$

by real coefficients $a_1, a_2, a_3, \cdots$ in the sense that

$$k\boldsymbol{H}k = a_1^2 + a_2^2 + a_3^2 + \cdots$$
$$k\boldsymbol{E}k = \eta_1a_1^2 + \eta_2a_2^2 + \eta_3a_3^2 + \cdots.$$

It is beyond our purpose to discuss the nature of the function $\underline{k}$ and of these characteristic functions in detail.

BRAUNSCHWEIG, GERMANY

A PRIORI LIMITATIONS FOR SOLUTIONS OF MONGE-AMPÈRE EQUATIONS. II*

BY
HANS LEWY

In this paper we are concerned with the convergence of solutions of elliptic and analytic Monge-Ampère equations. Theorem 1 gives the principal result of this paper. The example on p. 372 indicates the possibility of certain types of singularities which for linear elliptic equations cannot occur. Theorems 2 and 3 give sufficient conditions for the analyticity of the limit function. These conditions allow applications to certain problems of the differential geometry in the large. Our method consists in introducing a regularizing contact transformation which transforms convex functions into functions with bounded second derivatives and thus makes possible the reduction of Theorem 1 to the principal result of the first part of this paper.

1. **Regularizing contact transformation.** Consider the contact transformation T of an (x, y, z)-space into a (ξ, η, ζ)-space generated by the relation

$$0 = z + \zeta - x\xi - y\eta + \frac{x^2 + y^2}{2}.$$

It leads to the following transformation of a surface $z = z(x, y)$ with continuous first derivatives z_x and z_y

$$\begin{aligned} \xi &= x + z_x(x, y) \\ \eta &= y + z_y(x, y) \end{aligned} \tag{1}$$

$$\zeta = -z + x\xi + y\eta - (x^2 + y^2)/2. \tag{2}$$

Suppose $z(x, y)$ convex. Then (1) transforms an open rectangle R' of the (x, y)-plane into a domain of the (ξ, η)-plane in a one-to-one way.

For let (ξ_1, η_1) and (ξ_2, η_2) be the images of two distinct points (x_1, y_1) and (x_2, y_2), ϑ the angle between the vector joining (x_1, y_1) to (x_2, y_2) and the positive x-axis so that the distance between these two points is

$$(x_2 - x_1)\cos\vartheta + (y_2 - y_1)\sin\vartheta > 0.$$

As the derivative of z in the direction of the above vector is $z_x \cos\vartheta + z_y \sin\vartheta$, we conclude from the convexity of $z(x, y)$

* Presented to the Society, November 28, 1936; received by the editors September 1, 1936. Part I of this paper was published in these Transactions, vol. 37 (1935), pp. 417–434.

$$z_x(x_1, y_1) \cos \vartheta + z_y(x_1, y_1) \sin \vartheta \leqq z_x(x_2, y_2) \cos \vartheta + z_y(x_2, y_2) \sin \vartheta.$$

Hence

$$0 < (x_2 - x_1) \cos \vartheta + (y_2 - y_1) \sin \vartheta \leqq (\xi_2 - \xi_1) \cos \vartheta + (\eta_2 - \eta_1) \sin \vartheta$$

which implies

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 \leqq (\xi_2 - \xi_1)^2 + (\eta_2 - \eta_1)^2. \tag{3}$$

Using (3) for differentiation of (2) we obtain

$$x = \frac{\partial \zeta}{\partial \xi}, \qquad y = \frac{\partial \zeta}{\partial \eta} \tag{4}$$

which together with (3) shows that ζ has continuous derivatives in ξ and η.

If, moreover, z has continuous second derivatives with respect to (x, y), then ζ has continuous second derivatives in ξ and η all of which are numerically $\leqq 1$. For, by (1), the partial derivatives

$$\xi_x = 1 + z_{xx}, \quad \eta_x = \xi_y = z_{xy}, \quad \eta_y = 1 + z_{yy},$$

are continuous in (x, y). Hence the derivatives x_ξ, x_η, y_ξ, y_η of the inverse functions are continuous in (ξ, η) and we have†

$$\tag{5} \begin{aligned} \frac{\partial^2 \zeta}{\partial \xi^2} &= x_\xi = \frac{1 + z_{yy}}{(1 + z_{xx})(1 + z_{yy}) - z_{xy}^2}, \\ \frac{\partial^2 \zeta}{\partial \xi \partial \eta} &= \frac{- z_{xy}}{(1 + z_{xx})(1 + z_{yy}) - z_{xy}^2}, \cdots, \end{aligned}$$

$$\left|\frac{\partial^2 \zeta}{\partial \xi^2}\right|, \quad \left|\frac{\partial^2 \zeta}{\partial \xi \partial \eta}\right|, \quad \left|\frac{\partial^2 \zeta}{\partial \eta^2}\right| \leqq 1, \qquad \frac{\partial^2 \zeta}{\partial \xi^2} \frac{\partial^2 \zeta}{\partial \eta^2} - \left(\frac{\partial^2 \zeta}{\partial \xi \partial \eta}\right)^2 > 0. \tag{6}$$

Now consider a bounded sequence of continuously differentiable convex functions $z_n(x, y)$ defined in R'. Designate by T_n the transformation of a concentric open rectangle R, contained in R', into the (ξ, η)-plane as induced by T for the function $z_n(x, y)$, by T_n^{-1} the inverse of T_n. We proceed to show that there exists a subsequence of n such that

(i) $z_n(x, y)$ converge uniformly in R;

(ii) there exists an open set D in the (ξ, η)-plane contained in all $T_m(R)$ for m sufficiently large;

(iii) the inverse T_m^{-1} converges uniformly in D to a transformation T_∞^{-1};

(iv) $T_\infty^{-1}(D)$ contains a neighborhood of the center of R, and all those

† The denominator in (5) is >0 since the convexity of $z(x, y)$ implies $z_{xx} \geqq 0$, $z_{yy} \geqq 0$, $z_{xx}z_{yy} - z_{xy}^2 \geqq 0$.

points (ξ, η) which by T_∞^{-1} are mapped into the center of R have no limit point on the boundary of D.

The functions $z_n(x, y)$ are uniformly bounded and convex in R', hence equally continuous, so that (i) is satisfied and we have for $m>n$ in R

$$| z_m(x, y) - z_n(x, y) | < \epsilon_n$$

with $\epsilon_n \to 0$ as $n \to \infty$. If (x_0, y_0) is an arbitrary point of R, a circle γ_n about (x_0, y_0) with radius $\rho_n = 2(\epsilon_n)^{1/2}$ lies for sufficiently large n within R.

Denote by (ξ_n, η_n) the image of (x_0, y_0) in T_n, by ϑ the angle of the vector from (x_0, y_0) to a variable point (x_n', y_n') of the circumference of γ_n and the positive x-axis and by (ξ_m', η_m') the image of (x_n', y_n') in T_m. We have

$$\begin{aligned} z_{nx}(x_0, y_0) \cos\vartheta + z_{ny}(x_0, y_0) \sin\vartheta &\leqq \frac{z_n(x_n', y_n') - z_n(x_0, y_0)}{\rho_n} \\ &\leqq \frac{z_m(x_n', y_n') - z_m(x_0, y_0)}{\rho_n} + \frac{2\epsilon_n}{\rho_n} \\ &\leqq z_{mx}(x_n', y_n') \cos\vartheta + z_{my}(x_n', y_n') \sin\vartheta + \frac{2\epsilon_n}{\rho_n}. \end{aligned}$$

Hence

$$\xi_n \cos\vartheta + \eta_n \sin\vartheta < \xi_m' \cos\vartheta + \eta_m' \sin\vartheta.$$

Thus as ϑ varies from 0 to 2π the projection of the vector $(\xi_m' - \xi_n, \eta_m' - \eta_n)$ on the vector $(\cos\vartheta, \sin\vartheta)$ is positive or the angle between the two vectors is $<\pi/2$, thereby implying that the vector $(\xi_m' - \xi_n, \eta_m' - \eta_n)$ turns by 2π as the point (ξ_m', η_m') describes the image in T_m of the circumference of γ_n. Hence, for all $m>n$, (ξ_n, η_n) lies within $T_m(\gamma_n)$. Denoting by R_n the set of all those points of R whose distance from the boundary of R is $>\rho_n$, we may say that $T_m(R)$ contains $T_n(R_n)$ for $m>n$. Thus (ii) is proved with $D = T_n(R_n)$.

As the transformations T_m^{-1} are bounded and by (3) equicontinuous for all m we may pick out a suitable subsequence satisfying (iii).

Setting $(\xi_m, \eta_m) = T_m(x_0, y_0)$ and $(\xi_n', \eta_n') = T_n(x_n', y_n')$ for (x_n', y_n') on γ_n, we obtain in a way similar to the one used above

$$\xi_m \cos\vartheta + \eta_m \sin\vartheta < \xi_n' \cos\vartheta + \eta_n' \sin\vartheta$$

from which we conclude that for $m>n$ the point (ξ_m, η_m) lies in $T_n(\gamma_n)$. Now take n large enough so that there exists a point (x_0, y_0) of R_n at distance $>\rho_n$ from the boundary of R_n. As the corresponding points (ξ_m, η_m) lie all in $T_n(\gamma_n)$ there exists a converging subsequence tending to a point (ξ_0, η_0). In view of (3), the distance between $T_m^{-1}(\xi_m, \eta_m)$ and $T_m^{-1}(\xi_0, \eta_0)$ tends to zero if the dis-

tance between (ξ_m, η_m) and (ξ_0, η_0) tends to zero. Hence for our subsequence $T_m^{-1}(\xi_0, \eta_0)\to(x_0, y_0)$, or (x_0, y_0) is the image in T_∞^{-1} of a point (ξ_0, η_0) of $T_n(R_n)=D$. Furthermore, as T_m^{-1} maps $T_n(R_n)$ into a domain containing all points (x_0, y_0) of R_n at distance $>\rho_n$ from the boundary of R_n and as the mapping is one-to-one it follows that the boundary of $T_n(R_n)$ is mapped by T_m^{-1} into a set of points whose distance from the boundary of R_n is $\leqq\rho_n$. In view of the uniform convergence we conclude that T_∞^{-1} maps the boundary of $T_n(R_n)$ into a set of points distinct from the center of R. At the same time $T_\infty^{-1}T_n(R_n)$ contains a neighborhood of the center of R, namely, the set of points at distance $>\rho_n$ from the boundary of R_n, which proves (iv).

From (2) we derive by passing to the limit for $T_\infty^{-1}(\xi, \eta)=(x, y)$

$$\lim_{n\to\infty} \zeta_n(\xi, \eta) = \lim_{n\to\infty} (-z_n(x, y) + x\xi + y\eta - (x^2 + y^2)/2);$$

by (4) the derivatives of $\zeta(\xi, \eta)$ converge uniformly, hence evidently

$$\frac{\partial \zeta_n}{\partial \xi} \to \frac{\partial \lim \zeta_n}{\partial \xi}, \quad \frac{\partial \zeta_n}{\partial \eta} \to \frac{\partial \lim \zeta_n}{\partial \eta}.$$

If the limit function ζ is analytic in ξ and η, the map T_∞^{-1} is "schlicht" in the neighborhood of a point Q such that $T_\infty^{-1}(Q)$ is a point c near the center of R.

To prove this, let us assume the existence of another point Q' with $T_\infty^{-1}(Q')=c$. Call, for n in the above subsequence, σ_n the segment joining $T_n^{-1}(Q)$ to $T_n^{-1}(Q')$, and k_ρ a circle of small, arbitrary radius $\rho>0$ about Q and lying in D. The intersections of k_ρ with $T_n(\sigma_n)$ have at least one limit point L and we conclude from the uniform convergence of T_n^{-1} to T_∞^{-1} that $T_\infty^{-1}(L)=c$. Thus it is shown that if there are two points in D, Q and Q', such that $T_\infty^{-1}(Q)=T_\infty^{-1}(Q')=c$, then there exists an infinity of such points. Since the functions $x(\xi, \eta)$ and $y(\xi, \eta)$ determined by the transformation T_∞^{-1} are analytic, there exists a real curve C such that the two equations hold

$$(*) \qquad [x(\xi, \eta), y(\xi, \eta)] = c.$$

All points (ξ, η) such that (*) is satisfied lie within a closed set interior to D because T_∞^{-1} on the boundary of D differs from c as formerly proved. Taking local Puiseux developments it is easy to show that the above analytic curve has at each point a finite number of tangents and that no point of it is a "dead end," i.e., endpoint of one arc and of no other. Furthermore (*) implies that there are at most a finite number of singular points of C in D. Let us choose an arbitrary regular point on C and introduce an orientation on C in the neighborhood of this point. Continue the orientation on the same branch of the curve until we arrive at a singular point. Among the several

arcs ending at the singular point we pick one distinct from the one already oriented and continue the orientation on this arc. Proceeding this way we must finally encounter a point twice, as the length of our path is bounded on account of the finiteness of the number of singular points. Hence there exists a closed circuit δ consisting only of arcs of C, in particular one bounding a simply connected domain D'. Now $T_n^{-1}(\delta)$ is the boundary of a simply connected domain $T_n^{-1}(D')$ and $T_n^{-1}(\delta)\to c$. Hence $T_n^{-1}(D')\to c$, $T_\infty^{-1}(D')=c$ which is impossible since T_∞^{-1} is analytic and does not map identically into c.

2. We first prove the following theorem.

THEOREM 1. *Let A, B, C, E be analytic functions of five complex arguments u, v, x, p, q in a bounded and closed neighborhood N of a real point u_0, v_0, x_0, p_0, q_0 and suppose A, B, C, E depending on a parameter μ and converging uniformly in N as $\mu\to\mu^*$. Assume A, B, C, E real for real u, v, x, p, q,*

$$\Delta \equiv 4(AC+E) - B^2 > 0,$$

and $|\Delta|^{-1}$ uniformly bounded in N. Suppose that there exists a rectangle R with u_0, v_0 as center and such that there exists for $\mu\neq\mu^*$ a real and analytic function $x(u, v)$ which is a solution of the Monge-Ampère equation†

$$Ar + Bs + Ct + (rt - s^2) = E; \tag{7}$$

suppose that for this solution (u, v, x, p, q) remains in N, when (u, v) ranges over R. Then there exists a real neighborhood of (u_0, v_0) such that the solutions $x(u, v)$ corresponding to a subsequence of values of μ converge uniformly to a limit function $x^*(u, v)$ as $\mu\to\mu^*$. There exists an analytic relation between u, v, x^*, and $x^*(u, v)$ has continuous first derivatives. Moreover, $x^*(u, v)$ is analytic in u and v and satisfies the limit relation (7) for $\mu\to\mu^*$ at each point where a certain not identically vanishing analytic function $G(u, v)$ does not vanish.

Since

$$0 < \Delta = 4(A+t)(C+r) - (B-2s)^2,$$

we have $A+t$ and $C+r$ for each solution always $\neq 0$ and of the same sign. Hence in the sequence, there are infinitely many solutions for which $A+t$ is of the same sign, and we may even assume $A+t$ positive since in the opposite case we may replace A, B, C, E, and $x(u, v)$ by $-A$, $-B$, $-C$, $+E$, and $-x(u, v)$ without changing hypothesis nor conclusion of the theorem.

Let, for all μ of the subsequence,

$$|A|,\ |C|,\ \left|\frac{B}{2}\right| \leqq a.$$

† Here p, q, r, s, t denote, as usual, the partial derivatives of $x(u, v)$ of the first and second orders, respectively.

We find

$$\begin{aligned} 0 &< (r+C)(t+A) - \left(\frac{B}{2} - s\right)^2 \\ &\leqq (r+a)(t+a) - s^2 + \frac{B^2}{4} - B\left(\frac{B}{2} - s\right) \\ &\leqq (r+a)(t+a) - s^2 + a^2 + 2a\,\frac{r+a+t+a}{2} \\ &= (r+2a)(t+2a) - s^2. \end{aligned}$$

Hence $x(u, v)+a(u^2+v^2)$ is convex in R, and satisfies the transformed Monge-Ampère equation with the same discriminant Δ. It is therefore legitimate to *assume henceforth, that for all μ of a subsequence $x(u, v)$ is convex.* Let us then apply the contact transformation T of the (u, v, x)-space into a (ξ, η, ζ)-space. The Monge-Ampère equation thereby is transformed into another,

$$\tilde{A}\tilde{r} + \tilde{B}\tilde{s} + \tilde{C}\tilde{t} + \tilde{D}(\tilde{r}\tilde{t} - \tilde{s}^2) = \tilde{E}, \tag{8}$$

with

$$\tilde{A} = C - 1,\ \tilde{C} = A - 1,\ \tilde{B} = -B,\ \tilde{D} = 1 - E - A - C,\ \tilde{E} = -1. \tag{9}$$

The discriminant of the new equation equals that of the former:

$$4(\tilde{A}\tilde{C} + \tilde{D}\tilde{E}) - \tilde{B}^2 = 4(AC + E) - B^2 = \Delta.$$

The solution $\zeta(\xi, \eta)$ exists for all μ of a suitable subsequence in a common domain D of the (ξ, η)-plane and converges with its first derivatives to a limit function ζ^* and its derivatives. Let ξ_0, η_0 be an arbitrary, but fixed point in D, π_0 and κ_0 the derivatives of $\zeta^*(\xi, \eta)$ at (ξ_0, η_0). Replace $\zeta(\xi, \eta)$ by $\zeta(\xi, \eta)-\pi_0(\xi-\xi_0)-\kappa_0(\eta-\eta_0)=\bar{\zeta}(\xi, \eta)$. Then

$$\tilde{A}\frac{\partial^2\bar{\zeta}}{\partial\xi^2} + \tilde{B}\frac{\partial^2\bar{\zeta}}{\partial\xi\partial\eta} + \tilde{C}\frac{\partial^2\bar{\zeta}}{\partial\eta^2} + \tilde{D}\left(\frac{\partial^2\bar{\zeta}}{\partial\xi^2}\frac{\partial^2\bar{\zeta}}{\partial\eta^2} - \left(\frac{\partial^2\bar{\zeta}}{\partial\xi\partial\eta}\right)^2\right) = \tilde{E} \tag{10}$$

and

$$u = \frac{\partial\bar{\zeta}}{\partial\xi} + \pi_0, \qquad v = \frac{\partial\bar{\zeta}}{\partial\eta} + \kappa_0,$$

$$\begin{aligned} x = -\bar{\zeta} - \pi_0(\xi - \xi_0) - \kappa_0(\eta - \eta_0) + \xi\left(\frac{\partial\bar{\zeta}}{\partial\xi} + \pi_0\right) + \eta\left(\frac{\partial\bar{\zeta}}{\partial\eta} + \kappa_0\right) \\ - \frac{1}{2}\left[\left(\frac{\partial\bar{\zeta}}{\partial\xi} + \pi_0\right)^2 + \left(\frac{\partial\bar{\zeta}}{\partial\eta} + \kappa_0\right)^2\right], \end{aligned}$$

$$p = \xi - \frac{\partial \bar{\zeta}}{\partial \xi} - \pi_0, \qquad q = \eta - \frac{\partial \bar{\zeta}}{\partial \eta} - \kappa_0.$$

For all μ of the subsequence we may assume

$$\left(\frac{\partial \bar{\zeta}(\xi_0, \eta_0)}{\partial \xi}\right)^2 + \left(\frac{\partial \bar{\zeta}(\xi_0, \eta_0)}{\partial \eta}\right)^2 \leqq 1,$$

and our formulas show that the coefficients $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, $\tilde{E}$ considered as analytic functions of five independent variables $\xi, \eta, \bar{\zeta}, \partial\bar{\zeta}/\partial\xi, \partial\bar{\zeta}/\partial\eta$ (instead of u, v, x, p, q) exist in a complex neighborhood of ξ_0, η_0, $\bar{\zeta} = \zeta^*(\xi_0, \eta_0)$, $\partial\bar{\zeta}/\partial\xi = 0$, $\partial\bar{\zeta}/\partial\eta = 0$ and tend there uniformly to limit functions as $\mu \to \mu^*$.

By (6) we see that for all μ the second derivatives of our solution $\bar{\zeta}(\xi, \eta)$ are absolutely bounded by 1. Hence there exists a neighborhood of the point (ξ_0, η_0) of the (ξ, η)-plane in which, by the theorem on p. 432 of Part I of our paper, all $\bar{\zeta}(\xi, \eta)$ can be developed into power series in $\xi - \xi_0$ and $\eta - \eta_0$ such that there is a common majorant series for all μ which converges. Hence the limit $\zeta^*(\xi, \eta)$ of $\bar{\zeta}(\xi, \eta)$ is analytic in a neighborhood of ξ_0, η_0 and the same holds in the neighborhood of *any* interior point of D. As we have seen in §1 the transformation

$$u = \frac{\partial \zeta^*}{\partial \xi}, \qquad v = \frac{\partial \zeta^*}{\partial \eta}$$

maps the domain D into a set containing a neighborhood of the point (u_0, v_0) and there exists $\lim_{\mu \to \mu^*} x(u, v) = x^*(u, v) = -\zeta^*(\xi, \eta) + u\xi + \eta v - (u^2 + v^2)/2$. Thus it is shown that for a certain neighborhood of (u_0, v_0) the quantities u, v, x^* can be written as analytic functions of two variables ξ, η. Wherever the determinant $\partial(u, v)/\partial(\xi, \eta)$ differs from zero we may introduce u and v instead of ξ and η, and x^* becomes analytic in u and v. The relation $\partial(u, v)/\partial(\xi, \eta) = 0$ cannot be identically satisfied as the mapping $(\xi, \eta) \to (u, v)$ contains a neighborhood of (u_0, v_0). For all points with $\partial(u, v)/\partial(\xi, \eta) \neq 0$ we may differentiate with respect to u, v and find $\partial x^*(u, v)/\partial u = \xi - \partial\zeta^*/\partial\xi$, $\partial x^*/\partial v = \eta - \partial\zeta^*/\partial\eta$. Hence $\partial x^*/\partial u$ and $\partial x^*/\partial v$ are analytic functions of ξ, η and therefore continuous, even at points where $\partial(u, v)/\partial(\xi, \eta) = 0$ (since derivative at limit point equals limit of derivative). The proof of the theorem is now easily completed in view of the fact that the inverse transform of (8) is (7).

The following example shows that there are cases where there exists an analytic relation between u, v, x in a neighborhood of the origin, the function x can be considered as dependent on u and v and having continuous first derivatives there, and $x(u, v)$ is a solution of an analytic and elliptic Monge-

Ampère equation everywhere except on the u-axis where it fails to have second derivatives at all. The function is

$$x(u, v) = \frac{(3u)^{4/3}}{4} + \frac{v^2}{2},$$

the equation is

$$(-1 + p^2)r + rt - s^2 = 1.$$

3. There are wide classes of Monge-Ampère equations important in differential geometry such that the function $\partial(u, v)/\partial(\xi, \eta)$ cannot vanish at any interior point (u_0, v_0). In this case the limit function $x^*(u, v)$ is necessarily analytic.

Let us introduce two new variables α, β and solve the following initial problem for our variables ξ, η to be considered as functions of α and β:

$$\left(\frac{\partial\eta}{\partial\alpha} + i\frac{\partial\eta}{\partial\beta}\right)(\tilde{B} - 2\tilde{D}\tilde{s} + i\Delta^{1/2}) = \left(\frac{\partial\xi}{\partial\alpha} + i\frac{\partial\xi}{\partial\beta}\right)\cdot 2(\tilde{C} + \tilde{D}\tilde{r}), \tag{11}$$

$$\xi(\alpha, 0) = \alpha, \qquad \eta(\alpha, 0) = 0.$$

By the theorem of Cauchy-Kowalewski we have a solution $\xi(\alpha, \beta)$, $\eta(\alpha, \beta)$, analytic in the neighborhood of the origin. We have at the origin

$$\frac{\partial(\xi, \eta)}{\partial(\alpha, \beta)} = \frac{\partial\eta}{\partial\beta} = -2i(\tilde{C} + \tilde{D}\tilde{r})\frac{\left(1 + i\dfrac{\partial\xi}{\partial\beta}\right)}{(\tilde{B} - 2\tilde{D}\tilde{s} + i\Delta^{1/2})}$$

and hence may assert that there exists a neighborhood of the origin for which $\partial(\xi, \eta)/\partial(\alpha, \beta) \neq 0$ and the mapping $(\alpha, \beta) \to (\xi, \eta)$ is one-to-one. (11) may be written, with $=\partial/\partial\alpha + i\partial/\partial\beta$

$$(\tilde{B} + i\Delta^{1/2})\nabla\eta - 2\tilde{C}\nabla\xi - 2\tilde{D}\nabla\tilde{\zeta}_\xi = 0. \tag{12}$$

With the aid of $\tilde{\zeta}_\xi = u$, $\xi = u + p^*$, $\eta = v + q^*$ we obtain, in view of (9), for $\partial(u, v)/\partial(\xi, \eta) \neq 0$,

$$\begin{aligned}[(-B + i\Delta^{1/2})(1 + t^*) - 2(A - 1)s^*]\nabla v + [(-B + i\Delta^{1/2})s^* \\ - 2(A - 1)(r^* + 1) - 2(1 - E - A - C)]\nabla u = 0,\end{aligned}$$

and this is seen to be equivalent to

$$2(A + t^*)\nabla v + (-B + 2s^* - i\Delta^{1/2})\nabla u = 0 \tag{13'}$$

or

$$(-B - i\Delta^{1/2})\nabla u + 2A\nabla v + 2\nabla q^* = 0. \tag{13}$$

From (13′) we have

$$(-B + i\Delta^{1/2})\nabla v + 2C\nabla u + 2\nabla p^* = 0, \tag{14}$$

in view of

$$\Delta = 4(A + t^*)(C + r^*) - (B - 2s^*)^2.$$

Now (ξ, η) is analytic in (α, β), and u, v, p^*, q^*, x^* are analytic in (ξ, η), hence also in (α, β) and the relations (13), (14) hold independently of $\partial(u, v)/\partial(\xi, \eta) \neq 0$. We have moreover

$$\nabla x^* - p^*\nabla u - q^*\nabla v = 0. \tag{15}$$

(13), (14), (15) form the characteristic equations of (7) for the limit function $x^*(u, v)$. The mapping $(\alpha, \beta) \to (u, v)$ is one-to-one as both $(\alpha, \beta) \to (\xi, \eta)$ and $(\xi, \eta) \to (u, v)$ are one-to-one (see p. 368). We have

THEOREM 2. *If, in addition to the hypotheses of Theorem* 1, *we have* $\partial(u, v)/\partial(\alpha, \beta) \neq 0$, *then* $x^*(u, v)$ *is analytic in* (u, v).

The proof follows immediately from Theorem 1 and the relation

$$\frac{\partial(u, v)}{\partial(\alpha, \beta)} = \frac{\partial(u, v)}{\partial(\xi, \eta)} \; \frac{\partial(\xi, \eta)}{\partial(\alpha, \beta)}$$

which implies $\partial(u, v)/\partial(\alpha, \beta) = 0$ if $\partial(u, v)/\partial(\xi, \eta) = 0$.

We mention the following case which frequently presents itself in differential geometry: Let us operate with $\partial/\partial\alpha - i\partial/\partial\beta$ on (13) and (14) and take imaginary parts. On eliminating the derivatives of x^*, p^*, q^* with respect to (α, β) with the aid of (13), (14), (15), we obtain

$$\begin{aligned}
\frac{\partial^2 u}{\partial\alpha^2} + \frac{\partial^2 u}{\partial\beta^2} &= h_1\left[\left(\frac{\partial u}{\partial\alpha}\right)^2 + \left(\frac{\partial u}{\partial\beta}\right)^2\right] + h_2\left[\left(\frac{\partial v}{\partial\alpha}\right)^2 + \left(\frac{\partial v}{\partial\beta}\right)^2\right] \\
&\quad + h_3\left(\frac{\partial u}{\partial\alpha}\frac{\partial v}{\partial\alpha} + \frac{\partial u}{\partial\beta}\frac{\partial v}{\partial\beta}\right) + h_4\frac{\partial(u, v)}{\partial(\alpha, \beta)}, \\
\frac{\partial^2 v}{\partial\alpha^2} + \frac{\partial^2 v}{\partial\beta^2} &= h_1'\left[\left(\frac{\partial u}{\partial\alpha}\right)^2 + \left(\frac{\partial u}{\partial\beta}\right)^2\right] + \cdots + h_4'\frac{\partial(u, v)}{\partial(\alpha, \beta)}.
\end{aligned}$$

Now suppose that the coefficients $h_1, h_2, h_3, h_4, h_1', h_2', h_3', h_4'$ *depend only on* u *and* v. As they are analytic in (u, v), a result of ours† shows that then necessarily $\partial(u, v)/\partial(\alpha, \beta) \neq 0$.

† *On the non-vanishing of the Jacobian in certain one-to-one mappings*, Bulletin of the American Mathematical Society, vol. 42 (1936), pp. 689–692.

Consider the special case where the coefficients of the equation (7) depend only on (u, v). Then the elimination of x^*, p^*, q^* is obviously possible. Furthermore it becomes possible to widen the conditions of Theorem 1, and to state

THEOREM 2′. *Let A, B, C, E be analytic functions of the complex variables u, v for (u, v) in $N\,[\,|u-u_0| \leqq \epsilon,\ |v-v_0| \leqq \epsilon\,]$ and suppose A, B, C, E depending on μ and converging uniformly as $\mu \rightarrow \mu^*$. Assume A, B, C, E real for real u, v,*

$$\Delta \equiv 4(AC + E) - B^2 > 0,$$

and $|\Delta|^{-1}$ uniformly bounded for (u, v) in N. Suppose that there exists a real and analytic function $x(u, v)$ which is a solution of (7) as (u, v) ranges over the real rectangle $R\,[\,|u-u_0| < \epsilon,\ |v-v_0| < \epsilon\,]$, and that $x(u, v)$ is uniformly bounded as μ varies. Then there exists a subsequence of values μ such that the corresponding solutions $x(u, v)$ converge uniformly in every closed subregion of R to an analytic limit function $x^*(u, v)$, which is a solution of the limit equation (7).

Proceeding precisely as in the proof of Theorem 1, we find a suitable constant α such that $x(u, v) + a(u^2 + v^2)$ is convex for some subsequence of values μ, and, in view of the assumptions, bounded as (u, v) ranges over R. It follows that in every closed concentric rectangle contained in R the first derivatives of $x(u, v)$ are uniformly bounded as μ varies. Thus the assumptions of Theorem 2′ imply those of Theorem 1 and, on the other hand, also those of Theorem 2.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

APPLICATIONS OF THE THEORY OF BOOLEAN RINGS TO GENERAL TOPOLOGY*

BY

M. H. STONE

Introduction

In an earlier paper† we have developed an abstract theory of Boolean algebras and their representations by algebras of classes. We now relate this theory to the study of general topology.

The first part of our discussion is devoted to showing that the theory of Boolean rings is mathematically equivalent to the theory of locally-bicompact totally-disconnected topological spaces. In R we have already prepared the way for a topological treatment of the perfect representation of an arbitrary Boolean ring. Continuing in this way, we find that the perfect representation is converted by the introduction of a suitable topology into a space of the indicated type. We have no difficulty in inverting this result, proving that every locally-bicompact totally-disconnected topological space arises by the same procedure from a suitable Boolean ring. It is thus convenient to call the spaces corresponding in this manner to Boolean rings, Boolean spaces. The algebraic properties of Boolean rings can, of course, be correlated in detail with the topological properties of the corresponding Boolean spaces. A simple instance of the correlation is the theorem that the Boolean rings with unit are characterized as those for which the corresponding Boolean spaces are bicompact. A familiar example of a bicompact Boolean space is the Cantor discontinuum or ternary set, which we discuss at the close of Chapter I.

Having established this direct connection between Boolean rings and topology, we proceed in the second part of the discussion to considerations of a yet more general nature. We propose the problem of representing an arbitrary T_0-space by means of maps in bicompact Boolean spaces. Our solution of this problem embodies an explicit construction of such maps, which we shall now describe briefly. In a given T_0-space $\mathfrak{R}$, the open sets and the nowhere dense sets generate a Boolean ring, with $\mathfrak{R}$ as unit, which characterizes the topological structure of $\mathfrak{R}$. Those subrings which contain $\mathfrak{R}$ and which are so large that the interiors of their member sets constitute bases for $\mathfrak{R}$, also char-

* Presented to the Society (in part), February 25, 1933, and September 5, 1936; received by the editors June 1, 1936.

† Stone, these Transactions, vol. 40 (1936), pp. 37–111. In the sequel this paper will be designated by the letter R, citations being made in the form "R Theorem 10" or "R Definition 6."

acterize the structure of $\mathfrak{R}$. We are thus provided with an extensive family of Boolean rings which can be employed in the investigation of the given space $\mathfrak{R}$; we call them basic rings for $\mathfrak{R}$. In any basic ring for $\mathfrak{R}$, those sets with closure not containing a given point $\mathfrak{r}$ in $\mathfrak{R}$ constitute an ideal; and in the corresponding Boolean space, this ideal is represented by an open set with closed non-void complement $\mathfrak{X}(\mathfrak{r})$. The points of $\mathfrak{R}$ are thus represented by closed subsets of a related bicompact Boolean space. We find further that the topological structure of $\mathfrak{R}$ is characterized by the distribution of these representative closed sets. Thus we are able to reduce the study of general T_0-spaces to the examination of their maps in bicompact Boolean spaces. This reduction is not without advantage in the consideration of explicit topological problems, as we show in several illustrative applications.

The general mapping theory which we have outlined in the preceding paragraph is sufficiently complicated to suggest a search for simplifications. We turn naturally to the various strong separation properties such as regularity and normality. The investigation of the several possibilities which arise occupies the third chapter of our discussion. The general mapping theory, as previously developed, indicates the procedure for its own simplification and leads us at once to the consideration of a class of topological spaces to which little attention has been paid in the past. These spaces are characterized by the property that in them the regular open sets—that is, the interiors of closed sets—constitute bases. Since they are more general than the regular spaces, we call them semi-regular spaces. After discussing the semi-regular and regular spaces in detail, we consider the completely regular spaces. Here it is necessary for us to study the class of all bounded continuous real functions in a topological space. We obtain a reasonably complete algebraic insight into the structure of this class and its correlation with the structure of the underlying topological space. We are thereby enabled to complete the study of the maps of completely regular spaces. The normal spaces are, from the point of view adopted here, so special that we do not devote any separate consideration to them.

Plainly, we are engaged here in building a general abstract theory and must accordingly be occupied to a considerable extent with the elaboration of technical devices. Such preoccupation appears the more unavoidable, because the known instances of our theory are so special and so isolated that they throw little light upon the domain which we have wished to investigate. Nevertheless, we have not neglected to test our theory by applying it to specific problems of general set-theoretical topology, some of which have remained unsolved for several years. We may note in particular the propositions established in Theorems 52 and 53. While these applications lie in the domain

of set-theoretical topology, it is clear that the algebraic tendencies of our method relate it more nearly to recent developments in combinatorial topology. Indeed, it appears that by a process of gradual generalization combinatorial topologists have now arrived at a point of view very similar to that expounded here. Accordingly, we should expect that applications of the present theory could be made also to combinatorial problems. It seems clear, for instance, that the study of approximation by abstract complexes could conveniently be based upon the theory of Boolean maps, as developed here. We cannot, however, enter upon the discussion of such further applications at this time.

The present detailed exposition of our theory brings some corrections to previous summary announcements, as will be apparent from a reading of Chapter III, §1 below. The simplification of the mapping theory which was originally stated to be possible in general is now recognized as characteristic for the semi-regular spaces.* The general theory is accordingly somewhat more complicated than we originally supposed; but its applications, so far as we have examined them, require no essential modification. There are some applications described in the announcement just cited which we do not consider in the present paper. They are chiefly the ones dealing with dimension-theory. In Definitions 6 and 7 below we introduce the index of a map. The connections between the index and dimension-theory are pointed out in the indicated announcement, where references to the literature can be found. We hope to return to this subject at a later opportunity.

For general information concerning the elementary concepts of topology we refer the reader to the first two chapters of the recent book of Alexandroff and Hopf† and to a paper of Kuratowski‡ which deals with the algebraic properties of the closure operation. The latter paper will not be referred to explicitly below, since we shall give all the requisite algebraic reckonings in some detail. As we have already indicated, we shall confine our attention to the T_0-spaces. Our reason for doing so is that, among all the topological spaces which can be defined in terms of closure or of neighborhood systems, the T_0-spaces appear to be essentially the most general spaces which are of real interest. The T-spaces, as we have pointed out elsewhere,§ are more general than the T_0-spaces but only from a logical, rather than a topological, point of view: if one identifies the topologically indistinguishable points in a T-space, the result is a T_0-space.

* Note the oversimplifications in Theorems V_1, V_2, and V_3 in the Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 197–202.

† Alexandroff and Hopf, *Topologie* I, Leipzig, 1935, cited hereafter by the letters AH.

‡ Kuratowski, Fundamenta Mathematicae, vol. 3 (1922), pp. 182–199.

§ Stone, Matematicheskii Sbornik, vol. 1 (old series, vol. 43), No. 5 (1936–37).

The notations of the present paper are largely determined by the usages already established in R. Thus, for instance, in considering Boolean rings as abstract entities, we continue to use the symbols $\cdot$, $\vee$, $+$, $<$, $'$ corresponding to the symbols $\cdot$, $\cup$, Δ, $\subset$, $'$, of the theory of classes. In the algebra of the closure operation, we write A^- in place of $\overline{A}$ for convenience in putting down the more complicated expressions such as $A^{-\prime-\prime}$.

The contents of the present paper may be summarized systematically under the following headings: Chapter I, Boolean spaces: §1, Introduction of topological concepts; §2, Relations between algebra and topology; §3, Universal Boolean rings and spaces. Chapter II, Maps in Boolean spaces: §1, The general theory of maps; §2, Construction of Boolean maps; §3, Relation between algebraic and other maps; §4, Applications to the theory of extensions; §5, Totally-disconnected and discrete spaces. Chapter III, Stronger separation conditions: §1, Semi-regular spaces; §2, Regular spaces; §3, Completely regular spaces.

Chapter I. Boolean spaces

1. **Introduction of topological concepts.** We commence with the introduction of a topology in the abstract class in which the perfect representation of a Boolean ring has been constructed as described in R Chapter IV. It is natural to impose upon $\mathfrak{E}$, the class of all prime ideals in a Boolean ring A, a neighborhood topology based upon the special subclasses $\mathfrak{E}(\mathfrak{a})$, where $\mathfrak{a}$ is an ideal in A, or upon the still more restricted subclasses $\mathfrak{E}(a)=\mathfrak{E}(\mathfrak{a}(a))$, where a is an arbitrary element in A. The consequences of this procedure are stated in the following theorem.

Theorem 1. *Let A be a Boolean ring, $\mathfrak{a}$ an arbitrary ideal in A, $\mathfrak{E}$ the class of all prime ideals in A, $\mathfrak{E}(\mathfrak{a})$ the class of all prime ideals which are not divisors of $\mathfrak{a}$, and $\mathfrak{E}(a)$ the class $\mathfrak{E}(\mathfrak{a}(a))$ corresponding to the principal ideal $\mathfrak{a}=\mathfrak{a}(a)$. The topologies imposed upon $\mathfrak{E}$ through the introduction of the neighborhood systems* (1) *and* (2), *where*

(1) *each $\mathfrak{E}(\mathfrak{a})$ is assigned as a neighborhood of every element which it contains;*

(2) *each $\mathfrak{E}(a)$ is assigned as a neighborhood of every element which it contains;*

are equivalent. Under them, $\mathfrak{E}$ is a topological space with the properties:

(1) *$\mathfrak{E}$ is a totally-disconnected,* locally-bicompact H-space;*

* We use the term "totally-disconnected" to mean that, whenever $\mathfrak{p}$ and $\mathfrak{q}$ are distinct points of a topological space, there exist disjoint closed sets which contain $\mathfrak{p}$ and $\mathfrak{q}$ respectively and which have the entire space as their union. It is clear that a totally-disconnected topological space is necessarily an H-space.

(2) *the classes* $\mathfrak{E}(\mathfrak{a})$ *are characterized as the open sets in* $\mathfrak{E}$;

(3) *the classes* $\mathfrak{E}(a)$ *are characterized as the bicompact open sets in* $\mathfrak{E}$.

The character of $\mathfrak{E}$ *is equal to the cardinal number of* A *whenever* A *is infinite. The space* $\mathfrak{E}$ *is bicompact if and only if* A *has a unit* e.

The algebraic properties of the sets $\mathfrak{E}(\mathfrak{a})$ and $\mathfrak{E}(a)$ are sufficient to justify the introduction of the systems (1) and (2) as neighborhood-systems. In fact, the sets $\mathfrak{E}(\mathfrak{a})$ have the following properties: an arbitrary element $\mathfrak{p}$ in $\mathfrak{E}$, being a prime ideal in A, has as its neighborhoods the classes $\mathfrak{E}(\mathfrak{a})$ where $\mathfrak{p} \not\supset \mathfrak{a}$, and is contained in each of its neighborhoods; if $\mathfrak{E}(\mathfrak{a})$ and $\mathfrak{E}(\mathfrak{b})$ are neighborhoods of $\mathfrak{p}$, then their intersection $\mathfrak{E}(\mathfrak{a}\mathfrak{b}) = \mathfrak{E}(\mathfrak{a})\mathfrak{E}(\mathfrak{b})$ is also a neighborhood of $\mathfrak{p}$; and if $\mathfrak{E}(\mathfrak{a})$ contains $\mathfrak{p}$, then $\mathfrak{E}(\mathfrak{a})$ is a neighborhood of $\mathfrak{p}$. Thus the system (1) has the properties demanded of a neighborhood-system in a topological space;* under it $\mathfrak{E}$ becomes a topological space. Under the system (1), the sets $\mathfrak{E}(a)$ are special open sets. Since the classes $\mathfrak{E}(\mathfrak{a})$ are characterized algebraically as unions of classes $\mathfrak{E}(a)$ by virtue of the fact, established in R Theorem 67 (2), that the relations $\mathfrak{a} = \mathrm{S}_{a\epsilon\mathfrak{a}}\mathfrak{a}(a)$ and $\mathfrak{E}(\mathfrak{a}) = \sum_{a\epsilon\mathfrak{a}}\mathfrak{E}(a)$ are equivalent, we see that the classes $\mathfrak{E}(a)$ constitute a basis in the topological space $\mathfrak{E}$. In consequence, the assignment of these sets as neighborhoods in accordance with (2) provides a neighborhood-system equivalent to the system (1). Moreover, it is evident that the classes $\mathfrak{E}(\mathfrak{a})$ are characterized as the open subsets of $\mathfrak{E}$ under these equivalent neighborhood-systems. The system (1) is identified in this way as the absolute neighborhood-system.†

The nature of the topological space $\mathfrak{E}$ is now easily determined. To show that it is totally-disconnected, we start with distinct points $\mathfrak{p}$ and $\mathfrak{q}$ in $\mathfrak{E}$ and construct ideals $\mathfrak{a}$ and $\mathfrak{b}$ so that $\mathfrak{p}\epsilon\mathfrak{E}(\mathfrak{a})$, $\mathfrak{q}\epsilon\mathfrak{E}(\mathfrak{b})$, $\mathfrak{E}(\mathfrak{a})\mathfrak{E}(\mathfrak{b}) = 0$, $\mathfrak{E}(\mathfrak{a}) \cup \mathfrak{E}(\mathfrak{b}) = \mathfrak{E}$. Since $\mathfrak{p}$ and $\mathfrak{q}$ are distinct prime ideals in A, there exists an element a which belongs to $\mathfrak{q}$ but not to $\mathfrak{p}$. If we now put $\mathfrak{a} = \mathfrak{a}(a)$, $\mathfrak{b} = \mathfrak{a}'(a)$, we see that $\mathfrak{p} \not\supset \mathfrak{a}$, $\mathfrak{q} \not\supset \mathfrak{b}$, $\mathfrak{a}\mathfrak{b} = \mathfrak{o}$, $\mathfrak{a} \vee \mathfrak{b} = \mathfrak{e}$; and we can rewrite these relations in the desired form. Now the sets $\mathfrak{E}(\mathfrak{a})$, $\mathfrak{E}(\mathfrak{b})$ are open in $\mathfrak{E}$; since they are mutually complementary, they are also closed in $\mathfrak{E}$. Thus $\mathfrak{E}$ is seen to be totally-disconnected; and, in particular, to be an H-space. It remains for us to prove that $\mathfrak{E}$ is locally-bicompact. We shall do so by showing that every neighborhood $\mathfrak{E}(a)$ in the system (2) is a bicompact subspace of $\mathfrak{E}$. The open sets in the subspace $\mathfrak{E}(a)$ are precisely the sets $\mathfrak{E}(\mathfrak{a})\mathfrak{E}(a)$. The Heine-Borel-Lebesgue covering property, which is to be proved for $\mathfrak{E}(a)$, assumes the following form: if $\mathfrak{A}$ is any class of ideals $\mathfrak{a}$ such that $\sum_{\mathfrak{a}\epsilon\mathfrak{A}}\mathfrak{E}(\mathfrak{a})\mathfrak{E}(a) = \mathfrak{E}(a)$, then $\mathfrak{A}$ contains ideals $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ such that $\sum_{\nu=1}^{\nu=n}\mathfrak{E}(\mathfrak{a}_\nu)\mathfrak{E}(a) = \mathfrak{E}(a)$. The desired property is equiva-

* AH, p. 43, Satz IX.

† AH, p. 42.

lent to the following algebraic property: if $a \epsilon \mathrm{S}_{\mathfrak{a}\epsilon\mathfrak{A}}\mathfrak{a}$, then $a \epsilon \mathrm{S}_{\nu=1}^{\nu=n}\mathfrak{a}_\nu$ for appropriate ideals $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ in $\mathfrak{A}$. The latter property has been established in R Theorem 17. Conversely, we can show that an open subset $\mathfrak{E}(\mathfrak{a})$ of $\mathfrak{E}$ which, when considered as a subspace, is bicompact, is necessarily a set $\mathfrak{E}(a)$. The sets $\mathfrak{E}(a)$ contained in $\mathfrak{E}(\mathfrak{a})$ are open in this subspace; and they are precisely those sets corresponding to the elements a of the ideal $\mathfrak{a}$. The relation $\mathfrak{E}(\mathfrak{a}) = \sum_{a\epsilon\mathfrak{a}}\mathfrak{E}(a)$ yields elements $a_1, \cdots, a_n$ in $\mathfrak{a}$ such that $\mathfrak{E}(\mathfrak{a}) = \sum_{\nu=1}^{\nu=n}\mathfrak{E}(a_\nu)$ in accordance with the postulated Heine-Borel-Lebesgue covering property. The latter relation is equivalent to $\mathfrak{E}(\mathfrak{a}) = \mathfrak{E}(a)$ where $a = a_1 \vee \cdots \vee a_n$. This completes the topological characterization of the sets $\mathfrak{E}(a)$. It may be remarked that every $\mathfrak{E}(a)$, being bicompact, is necessarily a closed subset of $\mathfrak{E}$.

The character of the space $\mathfrak{E}$ can now be determined without difficulty. In case A is a finite ring, the space $\mathfrak{E}$ consists of n points where 2^n is the number of elements in A and $n \geqq 0$; every subset of $\mathfrak{E}$ is an open set and its points are all isolated. In case A has an infinite cardinal number $\mathfrak{c}$, the neighborhood-system (2) also has cardinal number $\mathfrak{c}$. Consequently the character of $\mathfrak{E}$ (that is, the least cardinal number belonging to a basis in $\mathfrak{E}$) cannot exceed $\mathfrak{c}$. On the other hand, if the open sets $\mathfrak{G}_\alpha$, where the index α runs over a fixed abstract class A, constitute a basis in $\mathfrak{E}$, every set $\mathfrak{E}(a)$ is the union of appropriate sets $\mathfrak{G}_\alpha$, $\alpha \epsilon \mathrm{A}$. The bicompactness of $\mathfrak{E}(a)$ thus yields indices $\alpha_1, \cdots, \alpha_n$ such that $\mathfrak{E}(a) = \sum_{\nu=1}^{\nu=n}\mathfrak{G}_{\alpha_\nu}$. Since every $\mathfrak{E}(a)$ is thus the union of a finite number of sets from the given basis, the cardinal number of the class of all sets $\mathfrak{E}(a)$ does not exceed that of the basis in question. We thus see that the character of $\mathfrak{E}$ is equal to $\mathfrak{c}$.

If A has a unit e, then $\mathfrak{E} = \mathfrak{E}(e)$ is bicompact in accordance with the preceding results. On the other hand, if $\mathfrak{E} = \mathfrak{E}(\mathfrak{e})$ is bicompact, then $\mathfrak{e}$ is a principal ideal, $\mathfrak{e} = \mathfrak{a}(a)$; and its generating element $e = a$ is a unit in A.

We proceed immediately to establish the converse of Theorem 1.

THEOREM 2. *If $\mathfrak{S}$ is a totally-disconnected locally-bicompact H-space, then the bicompact open subsets of $\mathfrak{S}$ constitute a Boolean ring A; and, if the class $\mathfrak{E}$ of prime ideals in A is topologized as in Theorem* 1, *then $\mathfrak{E}$ and $\mathfrak{S}$ are topologically equivalent.*

We shall first show that the subsets of $\mathfrak{S}$ which have been designated as members of A constitute a basis in $\mathfrak{S}$. If $\mathfrak{s}$ is any point in $\mathfrak{S}$ and $\mathfrak{G}$ any open set containing $\mathfrak{s}$, we have to establish the existence of a set $\mathfrak{H}$ in A such that $\mathfrak{s}\epsilon\mathfrak{H} \subset \mathfrak{G}$. Since $\mathfrak{S}$ is locally-bicompact, there exists an open set $\mathfrak{F}$ which satisfies the relation $\mathfrak{s}\epsilon\mathfrak{F} \subset \mathfrak{G}$ and which has closure $\mathfrak{F}^-$ with the property of bicompactness. If $\mathfrak{F} = \mathfrak{F}^-$, we may put $\mathfrak{H} = \mathfrak{F}$. Otherwise we take advantage of the total-disconnectedness of $\mathfrak{S}$: if $\mathfrak{t}$ is any point in $\mathfrak{F}^-\mathfrak{F}' \neq 0$, then $\mathfrak{s} \neq \mathfrak{t}$;

and there exists a set $\mathfrak{G}(\mathfrak{t})$ which contains $\mathfrak{t}$ but not $\mathfrak{s}$ and which, together with its complement, is both open and closed in $\mathfrak{S}$. Since $\mathfrak{F}^-$ is bicompact, the relation $\mathfrak{F}^- = \mathfrak{F}^-\mathfrak{F} \cup \sum_{\mathfrak{t}\in\mathfrak{F}^-\mathfrak{F}'} \mathfrak{F}^-\mathfrak{G}(\mathfrak{t})$, in which $\mathfrak{F}$ and $\mathfrak{G}(\mathfrak{t})$ are open sets, implies the existence of points $\mathfrak{t}_1, \cdots, \mathfrak{t}_n$ in $\mathfrak{F}^-\mathfrak{F}'$ such that $\mathfrak{F}^- = \mathfrak{F}^-\mathfrak{F} \cup \sum_{\nu=1}^{\nu=n} \mathfrak{F}^-\mathfrak{G}(\mathfrak{t}_\nu)$. We now put $\mathfrak{H} = \mathfrak{F}\mathfrak{G}'(\mathfrak{t}_1) \cdots \mathfrak{G}'(\mathfrak{t}_n) = \mathfrak{F}^-\mathfrak{G}'(\mathfrak{t}_1) \cdots \mathfrak{G}'(\mathfrak{t}_n)$. It is evident that $\mathfrak{H}$ is both open and closed in $\mathfrak{S}$, that it contains $\mathfrak{s}$, and that it is contained in both $\mathfrak{F}$ and $\mathfrak{G}$. Since $\mathfrak{H}$ is closed in $\mathfrak{S}$ and coincides with $\mathfrak{H}\mathfrak{F}^-$, it is also closed in $\mathfrak{F}^-$. Thus $\mathfrak{H}$ is bicompact when considered as a subspace of $\mathfrak{F}^-$; since the topology of $\mathfrak{H}$ is the same whether $\mathfrak{H}$ be regarded as a subspace of $\mathfrak{F}^-$ or of $\mathfrak{S}$, $\mathfrak{H}$ is bicompact also with respect to $\mathfrak{S}$. We thus complete the proof that $\mathfrak{S}$ has the desired basis.

We shall next show that the subsets of $\mathfrak{S}$ which are designated as members of A constitute a perfect algebra of classes in $\mathfrak{S}$, in accordance with R Definition 12. It is immediately evident that A contains $\mathfrak{A}\mathfrak{B}$ and $\mathfrak{A}\Delta\mathfrak{B}$ together with $\mathfrak{A}$ and $\mathfrak{B}$: for $\mathfrak{A}\mathfrak{B}$ is open since it is the intersection of open sets and bicompact since it is closed in the bicompact sets $\mathfrak{A}$ and $\mathfrak{B}$ respectively; the set $\mathfrak{A} \cup \mathfrak{B}$ is bicompact; and the set $\mathfrak{A}\Delta\mathfrak{B} = \mathfrak{A}\mathfrak{B}' \cup \mathfrak{A}'\mathfrak{B} \subset \mathfrak{A} \cup \mathfrak{B}$ is both open and closed in $\mathfrak{S}$ and is bicompact since it is closed in the bicompact space $\mathfrak{A} \cup \mathfrak{B}$. Thus A is a Boolean ring with classes as elements. Since A is a basis in the H-space $\mathfrak{S}$, it is a reduced algebra of classes in accordance with R Definition 10: each point of $\mathfrak{S}$ is contained in some member of A and is the sole point common to all such members of A. In order to show that A is perfect, we prove that $\sum_{\mathfrak{A}\in\mathfrak{a}} \mathfrak{A} = \sum_{\mathfrak{B}\in\mathfrak{b}} \mathfrak{B}$ implies $\mathfrak{a} = \mathfrak{b}$ whenever $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in A. It is evidently sufficient to prove that $\sum_{\mathfrak{A}\in\mathfrak{a}} \mathfrak{A} \subset \sum_{\mathfrak{B}\in\mathfrak{b}} \mathfrak{B}$ implies $\mathfrak{a} \subset \mathfrak{b}$. If $\mathfrak{A}$ is an arbitrary element of $\mathfrak{a}$, we have $\mathfrak{A} \subset \sum_{\mathfrak{B}\in\mathfrak{b}} \mathfrak{B}$ by hypothesis. Hence we can write $\mathfrak{A} = \sum_{\mathfrak{B}\in\mathfrak{b}} \mathfrak{A}\mathfrak{B}$ and even, by virtue of the bicompactness of $\mathfrak{A}$, the stronger relation $\mathfrak{A} = \sum_{\nu=1}^{\nu=n} \mathfrak{A}\mathfrak{B}_\nu$ where $\mathfrak{B}_1, \cdots, \mathfrak{B}_n$ are in $\mathfrak{b}$. Since the latter relation implies $\mathfrak{A}\in\mathfrak{b}$, we conclude that $\mathfrak{a} \subset \mathfrak{b}$. Thus A is perfect in accordance with R Definition 12.

Now by R Theorem 69, the algebra A is equivalent to its perfect representation $B(A)$: in other words, there exists a biunivocal correspondence between $\mathfrak{S}$ and $\mathfrak{E} = \mathfrak{E}(A)$ which carries the classes $\mathfrak{A}$ belonging to A into their respective representative classes $\mathfrak{E}(\mathfrak{A})$ in a biunivocal manner. Since A is a basis in $\mathfrak{S}$ and the classes $\mathfrak{E}(\mathfrak{A})$ constitute a basis in $\mathfrak{E}$ by Theorem 1, we conclude that $\mathfrak{S}$ and $\mathfrak{E}$ are topologically equivalent.

The correspondence between Boolean rings and topological spaces established in Theorems 1 and 2 justifies the introduction of the following definitions and notations.

Definition 1. *A totally-disconnected locally-bicompact H-space is called a Boolean space.*

Definition 2. *A Boolean space $\mathfrak{B}$ is called a representative of a Boolean ring A, symbolically, $\mathfrak{B}=\mathfrak{B}(A)$, if $\mathfrak{B}$ is topologically equivalent to the particular Boolean space $\mathfrak{C}=\mathfrak{C}(A)$ described in Theorem* 1.

Before developing further details of the relations connecting Boolean rings and their representative Boolean spaces, we note a few useful topological facts.

Theorem 3. *The closed subsets and the open subsets of a Boolean space are Boolean spaces. A continuous image of a bicompact Boolean space is necessarily a bicompact topological space; it is therefore a Boolean space if and only if it is totally-disconnected.*

Let $\mathfrak{F}$ be a closed subset of a Boolean space $\mathfrak{B}$. If $\mathfrak{x}_1$ and $\mathfrak{x}_2$ are distinct points in $\mathfrak{F}$, then there exist closed sets $\mathfrak{F}_1$ and $\mathfrak{F}_2$ such that $\mathfrak{x}_1\epsilon\mathfrak{F}_1$, $\mathfrak{x}_2\epsilon\mathfrak{F}_2$, $\mathfrak{F}_1\mathfrak{F}_2=0$, $\mathfrak{F}_1 \cup \mathfrak{F}_2=\mathfrak{B}$. The sets $\mathfrak{H}_1=\mathfrak{F}\mathfrak{F}_1$, $\mathfrak{H}_2=\mathfrak{F}\mathfrak{F}_2$ are closed in $\mathfrak{F}$ and have the properties $\mathfrak{x}_1\epsilon\mathfrak{H}_1$, $\mathfrak{x}_2\epsilon\mathfrak{H}_2$, $\mathfrak{H}_1\mathfrak{H}_2=0$, $\mathfrak{H}_1 \cup \mathfrak{H}_2=\mathfrak{F}$. Hence the space $\mathfrak{F}$ is totally-disconnected; in particular $\mathfrak{F}$ is an H-space. If $\mathfrak{H}$ is any open subset of $\mathfrak{F}$ and $\mathfrak{x}$ is any point in $\mathfrak{H}$, then $\mathfrak{H}=\mathfrak{G}\mathfrak{F}$ where $\mathfrak{G}$ is open in $\mathfrak{B}$. Since $\mathfrak{G}$ contains $\mathfrak{x}$, there exists an open set $\mathfrak{G}_1$ which has the property $\mathfrak{x}\epsilon\mathfrak{G}_1 \subset \mathfrak{G}$ and possesses a bicompact closure $\mathfrak{G}_1^-$ in $\mathfrak{B}$. The set $\mathfrak{H}_1=\mathfrak{G}_1\mathfrak{F}$ is open in $\mathfrak{F}$ and has the property $\mathfrak{x}\epsilon\mathfrak{H}_1 \subset \mathfrak{H}$. Furthermore its closure in $\mathfrak{F}$ is the closed set $(\mathfrak{G}_1\mathfrak{F})^-\mathfrak{F} \subset \mathfrak{G}_1^-\mathfrak{F} \subset \mathfrak{G}_1^-$; being closed also in $\mathfrak{G}_1^-$, it is bicompact. Thus $\mathfrak{F}$ has all the properties of a Boolean space.

Let $\mathfrak{G}$ be an open subset of a Boolean space $\mathfrak{B}$. Then $\mathfrak{G}$ is a totally-disconnected H-space by the same reasoning as that applied to $\mathfrak{F}$ in the preceding paragraph. The proof that $\mathfrak{G}$ is locally-bicompact proceeds as follows: the open sets in the subspace $\mathfrak{G}$ are precisely the open sets of $\mathfrak{B}$ which are contained in $\mathfrak{G}$; if $\mathfrak{x}$ is any point of $\mathfrak{G}$ and $\mathfrak{G}_1$ an open set such that $\mathfrak{x}\epsilon\mathfrak{G}_1 \subset \mathfrak{G}$, then there exists an open set $\mathfrak{H}$ which has the property $\mathfrak{x}\epsilon\mathfrak{H} \subset \mathfrak{G}_1 \subset \mathfrak{G}$ and possesses a bicompact closure $\mathfrak{H}^-=\mathfrak{H}$; and the closure of $\mathfrak{H}$ in the space $\mathfrak{G}$ is given by $\mathfrak{H}^-\mathfrak{G}=\mathfrak{H}\mathfrak{G}=\mathfrak{H}$.

It is well known that every continuous image of a bicompact H-space is a bicompact topological space.* The final statement of the theorem is thus obvious.

A proof of Theorem 3 can also be obtained in an indirect way from the correspondence between Boolean rings and Boolean spaces, as will be seen in the following section; but it is preferable to have the direct proof before us in stating our results.

2. **Relations between algebra and topology.** We shall now develop in

* AH, p. 95, Satz 1.

much greater detail the consequences of the relations between algebra and topology established in §1. We have the following general theorem.

THEOREM 4. *The algebraic theory of Boolean rings is mathematically equivalent to the topological theory of Boolean spaces by virtue of the following relations:*

(1) *every Boolean ring has a representative Boolean space; every Boolean space is the representative of some Boolean ring; and two Boolean rings are isomorphic if and only if their representatives are topologically equivalent;*

(2) *the group of automorphisms of a Boolean ring is isomorphic to the topological group of an arbitrary representative of the ring;*

(3) *the representatives of Boolean rings which are isomorphic to the various ideals in a Boolean ring A are characterized topologically as the open subsets of an arbitrary representative of A; in particular, $\mathfrak{E}(\mathfrak{a})$ is a representative of the ideal $\mathfrak{a}$ in A;*

(4) *the representatives of the homomorphs of a Boolean ring A are characterized topologically as the closed subsets of an arbitrary representative of A; in particular, $\mathfrak{E}'(\mathfrak{a})$ is a representative of the quotient ring $A/\mathfrak{a}$;*

(5) *the representatives of Boolean rings with unit are characterized topologically by the property of bicompactness.*

This theorem merely collects in a new form results already proved in R and in Theorems 1 and 2 above. To establish (1) we merely compare R Theorem 69, Theorem 1, and Theorem 2. We then see that (2) follows directly from (1). We have already established (5) as part of Theorem 1. Using (1) in conjunction with Theorem 1 (2), we see that (3) and (4) are established as soon as we can prove the special assertions to the effect that $\mathfrak{E}(\mathfrak{a})$ and $\mathfrak{E}'(\mathfrak{a})$ are representatives of $\mathfrak{a}$ and $A/\mathfrak{a}$ respectively. To show that $\mathfrak{E}(\mathfrak{a})$ is a representative of $\mathfrak{a}$, considered as a Boolean ring, we need only show that the sets $\mathfrak{E}(a)$, $a\epsilon\mathfrak{a}$, constitute a perfect algebra of classes in $\mathfrak{E}(\mathfrak{a})$ isomorphic to the ring $\mathfrak{a}$. Now the ideals in the ring $\mathfrak{a}$ are precisely the ideals $\mathfrak{b}$ in A such that $\mathfrak{b}\subset\mathfrak{a}$; and the relation $\mathfrak{b}\subset\mathfrak{a}$ is equivalent by R Theorem 67 to the relation $\mathfrak{E}(\mathfrak{b})\subset\mathfrak{E}(\mathfrak{a})$. That theorem shows further that the sets $\mathfrak{E}(\mathfrak{b})$ contained in $\mathfrak{E}(\mathfrak{a})$ constitute an algebra of classes isomorphic to the algebra of all ideals in $\mathfrak{a}$ in the precise sense stated there. On specializing to the sets $\mathfrak{E}(a)$ corresponding to the principal ideals $\mathfrak{b}=\mathfrak{a}(a)$, $a\epsilon\mathfrak{a}$, in the ring $\mathfrak{a}$, we see that these sets constitute an algebra of classes which is reduced and perfect in accordance with R Definitions 10 and 12. It should be noted that the sets $\mathfrak{E}(\mathfrak{b})$ contained in $\mathfrak{E}(\mathfrak{a})$ are precisely the open sets in the subspace $\mathfrak{E}(\mathfrak{a})$ of $\mathfrak{E}$: for every such open set is expressible, in harmony with Theorem 1 and R Theorem 67, in the form $\mathfrak{E}(\mathfrak{b})=\mathfrak{E}(\mathfrak{c})\mathfrak{E}(\mathfrak{a})$, where $\mathfrak{c}$ is an arbitrary ideal in A

and $\mathfrak{b}=\mathfrak{c}\mathfrak{a}$ is an ideal contained in $\mathfrak{a}$. The proof that $\mathfrak{E}'(\mathfrak{a})$ is a representative of $A/\mathfrak{a}$ is similar. We know from R Theorem 68 that the sets $\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ constitute a reduced perfect algebra of classes in $\mathfrak{E}'(\mathfrak{a})$ isomorphic to $A/\mathfrak{a}$; and that the sets $\mathfrak{E}(\mathfrak{b})\mathfrak{E}'(\mathfrak{a})$ constitute an algebra of classes isomorphic to the algebra of the ideals in $A/\mathfrak{a}$. The latter sets are precisely the open sets in the space $\mathfrak{E}'(\mathfrak{a})$. The desired result is thus established. We note that the relations between the spaces $\mathfrak{E}(\mathfrak{a})$, $\mathfrak{E}'(\mathfrak{a})$ and the Boolean rings $\mathfrak{a}$, $A/\mathfrak{a}$ imply that these spaces are Boolean spaces. Combining this fact with Theorem 1 (2), we obtain an indirect proof of the first part of Theorem 3.

It is a matter of some interest to characterize topologically the various classes of ideals introduced in R Definition 8. We find the following result, which is most conveniently stated in terms of the perfect representation $\mathfrak{E}(A)$.

THEOREM 5. *If $\mathfrak{a}$ is an ideal in a Boolean ring A, then $\mathfrak{E}(\mathfrak{a}')$ coincides with the exterior of $\mathfrak{E}(\mathfrak{a})$. Hence the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to ideals in the respective classes $\mathfrak{J}$, $\mathfrak{N}$, $\mathfrak{S}$, $\mathfrak{P}^*$, $\mathfrak{P}$ of* R *Definition* 8 *are characterized topologically as follows*:

(1) *the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to arbitrary ideals are the open sets in $\mathfrak{E}$;*
(2) *the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to normal ideals are the regular open sets in $\mathfrak{E}$;*
(3) *the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to simple ideals are the open-and-closed sets in $\mathfrak{E}$, or, equivalently, the sets in $\mathfrak{E}$ with void boundaries;*
(4) *the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to semiprincipal ideals are the open sets in $\mathfrak{E}$ which are bicompact or are complements of bicompact open sets;*
(5) *the sets $\mathfrak{E}(\mathfrak{a})$ corresponding to principal ideals are the bicompact open sets in $\mathfrak{E}$.*

We take (1) directly from Theorem 1 (2). It is then obvious from R Definition 7 and R Theorem 19 that $\mathfrak{E}(\mathfrak{a}')$ is the maximal open set disjoint from $\mathfrak{E}(\mathfrak{a})$; in other words, $\mathfrak{E}(\mathfrak{a}')$ coincides with $\mathfrak{E}^{-\prime}(\mathfrak{a})$, the exterior of $\mathfrak{E}(\mathfrak{a})$. Since $\mathfrak{a}''=\mathfrak{a}$ if and only if $\mathfrak{E}(\mathfrak{a}'')=\mathfrak{E}(\mathfrak{a})$, we see that the normal ideals are characterized by the relation $\mathfrak{E}(\mathfrak{a})=\mathfrak{E}^{-\prime-\prime}(\mathfrak{a})$, which identifies their representative open sets as the regular open sets. The simple ideals $\mathfrak{a}$ are those for which $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$, or, equivalently, $\mathfrak{E}(\mathfrak{a})\cup\mathfrak{E}(\mathfrak{a}')=\mathfrak{E}$. Hence they are characterized by the property $\mathfrak{E}'(\mathfrak{a})=\mathfrak{E}(\mathfrak{a}')$ or $\mathfrak{E}(\mathfrak{a})=\mathfrak{E}^{-}(\mathfrak{a})$. Their representative sets are thus characterized by the property of being closed as well as open. Now in any topological space, the relations $\mathfrak{H}=\mathfrak{H}^-$, $\mathfrak{H}'=\mathfrak{H}'^-$, which express the fact that $\mathfrak{H}$ is both closed and open, are together equivalent to the relation $\mathfrak{H}^-\mathfrak{H}'^-=0$ which expresses the fact that the boundary of $\mathfrak{H}$ is void: for, when the first relations hold, we have $\mathfrak{H}^-\mathfrak{H}'^-=\mathfrak{H}\mathfrak{H}'=0$; and, when the second relation holds, we have $\mathfrak{H}\subset\mathfrak{H}^-$, $\mathfrak{H}'\subset\mathfrak{H}'^-$, $\mathfrak{H}^-\mathfrak{H}'^-=0$, and hence $\mathfrak{H}=\mathfrak{H}^-$, $\mathfrak{H}'=\mathfrak{H}'^-$. Thus the simple ideals $\mathfrak{a}$ are characterized by the fact that the

corresponding sets $\mathfrak{E}(\mathfrak{a})$ have void boundaries, as stated in (3). We pass next to the consideration of (5). Here we have merely repeated the characterization already given in Theorem 1. Then by combining (3) and (5) we obtain (4), since an ideal $\mathfrak{a}$ is semiprincipal if and only if both ideals $\mathfrak{a}$ and $\mathfrak{a}'$ are simple and at least one of them principal.

By reference to R Theorems 38, 59, and 67, we see immediately that Theorem 5 leads to the following characterization of the sets corresponding to prime ideals:

THEOREM 6. *The set $\mathfrak{E}(\mathfrak{p})$ corresponding to an arbitrary prime ideal $\mathfrak{p}$ in a Boolean ring A is the open set $\mathfrak{E}-\{\mathfrak{p}\}$; the ideal $\mathfrak{p}$ is normal, and hence semiprincipal, if and only if $\mathfrak{p}$ is an isolated point in the Boolean space $\mathfrak{E}=\mathfrak{E}(A)$.*

Using Theorem 5 (3), we now obtain an important result.

THEOREM 7. *The representatives of Boolean rings B with the properties*

(1) *B is isomorphic to a subring $\mathfrak{B}$ of the ring $\mathfrak{S}$ of all simple ideals in a Boolean ring A;*

(2) *if $\mathfrak{C}$ is an ideal in $\mathfrak{B}$, the relations $\mathrm{S}_{\mathfrak{a}\in\mathfrak{C}}\mathfrak{a}=\mathfrak{e}$ and $\mathfrak{C}=\mathfrak{B}$ are equivalent;*

are characterized topologically as those Boolean spaces which are continuous images of an arbitrary representative of A. In particular, the representatives of Boolean rings B with the following properties:

(1) *B is isomorphic to a subring $\mathfrak{b}$ of a Boolean ring A with unit e;*

(2) *e is an element of $\mathfrak{b}$;*

are characterized topologically as the totally-disconnected continuous images of an arbitrary representative of A.

By Theorem 4 we may restrict our attention to the subrings $\mathfrak{B}$ and the space $\mathfrak{E}=\mathfrak{E}(A)$ in establishing the first part of the present theorem. Let $\mathfrak{A}$ be a Boolean space which is a continuous image of $\mathfrak{E}(A)$ by virtue of a correspondence $\mathfrak{s}=f(\mathfrak{p})$. Then there exists a Boolean ring B such that $\mathfrak{A}=\mathfrak{A}(B)$; and, if $\mathfrak{F}$ is any bicompact open set in $\mathfrak{A}$, its antecedent $f^{-1}(\mathfrak{F})$ is both open and closed in $\mathfrak{E}$. By Theorem 5 (3) there exists a simple ideal in A such that $f^{-1}(\mathfrak{F})=\mathfrak{E}(\mathfrak{a})$. The algebras with the classes $\mathfrak{F}$, $f^{-1}(\mathfrak{F})$ as elements are isomorphic to each other and to B, while the associated ideals $\mathfrak{a}$ constitute a subring of $\mathfrak{S}$ also isomorphic to B. This subring we designate as $\mathfrak{B}$. We can now write $\mathfrak{A}=\mathfrak{A}(\mathfrak{B})$ by virtue of the isomorphism between B and $\mathfrak{B}$. If $\mathfrak{C}$ is any ideal in $\mathfrak{B}$, it is represented by an open set $\mathfrak{A}(\mathfrak{C})$ in $\mathfrak{A}$. It is evident that $f^{-1}(\mathfrak{A}(\mathfrak{C}))=\sum_{\mathfrak{a}\in\mathfrak{C}}\mathfrak{E}(\mathfrak{a})=\mathfrak{E}(\mathrm{S}_{\mathfrak{a}\in\mathfrak{C}}\mathfrak{a})$ and hence that the relations $\mathfrak{C}=\mathfrak{B}$, $\mathfrak{A}(\mathfrak{C})=\mathfrak{A}$, and $\mathrm{S}_{\mathfrak{a}\in\mathfrak{C}}\mathfrak{a}=\mathfrak{e}$ are equivalent. On the other hand, let $\mathfrak{B}$ be a Boolean ring with the properties (1) and (2). The elements $\mathfrak{a}$ of $\mathfrak{B}$ can then be represented in two ways: as the bicompact open subsets $\mathfrak{A}(\mathfrak{a})$ of a representative $\mathfrak{A}=\mathfrak{A}(\mathfrak{B})$, and as certain open-and-closed sets $\mathfrak{E}(\mathfrak{a})$, $\mathfrak{a}\in\mathfrak{B}\subset\mathfrak{S}$, in $\mathfrak{E}=\mathfrak{E}(A)$. The corre-

spondence $\mathfrak{E}(\mathfrak{a}) \longleftrightarrow \mathfrak{A}(\mathfrak{a})$, $\mathfrak{a} \epsilon \mathfrak{B}$, then induces a point correspondence between $\mathfrak{E}$ and $\mathfrak{A}$. If $(\mathfrak{s}, \mathfrak{p})$ is a pair of points $\mathfrak{s}$ and $\mathfrak{p}$ in $\mathfrak{A}$ and $\mathfrak{E}$ respectively such that the relations $\mathfrak{A}(\mathfrak{a}) \subset \mathfrak{A} - \{\mathfrak{s}\}$ and $\mathfrak{E}(\mathfrak{a}) \subset \mathfrak{E} - \{\mathfrak{p}\}$ are either *both* true or *both* false whatever the element $\mathfrak{a}$ in $\mathfrak{B}$, we set $\mathfrak{s} = f(\mathfrak{p})$. We see that every $\mathfrak{s}$ in $\mathfrak{A}$ belongs to at least one such pair: for the relation $\mathfrak{A}(\mathfrak{a}) \subset \mathfrak{A} - \{\mathfrak{s}\}$ determines a prime ideal $\mathfrak{P}$ in $\mathfrak{B}$; and, by hypothesis, the relation $\mathfrak{P} \neq \mathfrak{B}$ implies $\sum_{\mathfrak{a} \epsilon \mathfrak{P}} \mathfrak{E}(\mathfrak{a}) = \mathfrak{E}(\mathrm{S}_{\mathfrak{a} \epsilon \mathfrak{P}} \mathfrak{a}) \neq \mathfrak{E}(\mathfrak{e}) = \mathfrak{E}$. Furthermore we see that each $\mathfrak{p}$ belongs to exactly one such pair: for those elements $\mathfrak{a}$ in $\mathfrak{B}$ which satisfy the equivalent relations $\mathfrak{a} \subset \mathfrak{p}$, $\mathfrak{E}(\mathfrak{a}) \subset \mathfrak{E}(\mathfrak{p})$ constitute a prime ideal $\mathfrak{P}$ in $\mathfrak{B}$ in accordance with R Theorems 36 and 41; and there is then exactly one point $\mathfrak{s}$ in $\mathfrak{A}$ such that $\sum_{\mathfrak{a} \epsilon \mathfrak{P}} \mathfrak{A}(\mathfrak{a}) \subset \mathfrak{A} - \{\mathfrak{s}\}$. The relation $\mathfrak{s} = f(\mathfrak{p})$ thus sets $\mathfrak{A}$ in univocal correspondence with $\mathfrak{E}$. Since the sets $\mathfrak{A}(\mathfrak{a})$ constitute a basis in $\mathfrak{A}$ and the sets $\mathfrak{E}(\mathfrak{a})$ are open in $\mathfrak{E}$, it is sufficient in proving that $f(\mathfrak{p})$ is continuous, to show that $f^{-1}(\mathfrak{A}(\mathfrak{a})) = \mathfrak{E}(\mathfrak{a})$ for every $\mathfrak{a}$ in $\mathfrak{B}$. If $\mathfrak{a}$ is fixed, the points $\mathfrak{s}$ and $\mathfrak{p}$ specified by the relations $f(\mathfrak{p}) = \mathfrak{s} \epsilon \mathfrak{A}(\mathfrak{a})$ are by definition precisely those for which the relations $\mathfrak{A}(\mathfrak{a}) \subset \mathfrak{A} - \{\mathfrak{s}\}$, $\mathfrak{E}(\mathfrak{a}) \subset \mathfrak{E} - \{\mathfrak{p}\}$ are both false; and the desired relation $f^{-1}(\mathfrak{A}(\mathfrak{a})) = \mathfrak{E}(\mathfrak{a})$ is thus established.

The second part of the theorem follows by specialization of what has just been proved. If A has a unit e, then $\mathfrak{S}$ is isomorphic to A by R Theorems 25, 30, and 31. Hence the rings $\mathfrak{B}$ with the properties (1) and (2) of the first part of the theorem can be replaced by the subrings $\mathfrak{b}$ of A which contain e: for every subring of $\mathfrak{S}$ is isomorphic to a subring of A; and property (2), when expressed in terms of the generating elements of the principal ideals involved, asserts the equivalence of the relations $\mathrm{S}_{a \epsilon \mathfrak{c}} a = e$ and $\mathfrak{c} = \mathfrak{b}$ for an arbitrary ideal $\mathfrak{c}$ in $\mathfrak{b}$, and therefore degenerates, in accordance with R Theorem 17, into the simpler condition $e \epsilon \mathfrak{b}$. Theorems 1 and 3 now permit us to assert the second part of the present theorem.

As a final illustration of the connections between algebra and topology developed in the preceding theorems, we may comment informally on two distinct algebraic problems. First let us consider the determination of the ideal product of a non-void class $\mathfrak{A}$ of prime ideals $\mathfrak{p}$ in a Boolean ring A. It is clear that the complete solution of this problem can be given in the following topological terms: if $\mathfrak{a} = \prod_{\mathfrak{p} \epsilon \mathfrak{A}} \mathfrak{p}$, then $\mathfrak{E}(\mathfrak{a}) = [\prod_{\mathfrak{p} \epsilon \mathfrak{A}} \mathfrak{E}(\mathfrak{p})]'^{-\prime} = [\sum_{\mathfrak{p} \epsilon \mathfrak{A}} \{\mathfrak{p}\}]^{-\prime}$. Then let us consider the problem of representative elements in the modular classes of a Boolean ring A, recently discussed by v. Neumann and the writer.* In topological terms, this problem can now be rephrased as follows: if $\mathfrak{F}$ is any closed subset of a Boolean space $\mathfrak{B}$, it is required to determine in each class of sets intersecting $\mathfrak{F}$ in a fixed bicompact open subset of that

* J. v. Neumann and M. H. Stone, Fundamenta Mathematicae, vol. 25 (1935), pp. 353–378.

subspace a representative member $\mathfrak{G}$ so that (1) $\mathfrak{G}$ is a bicompact open subset of $\mathfrak{B}$, (2) the Boolean rings generated by the sets $\mathfrak{G}$ and $\mathfrak{G}\mathfrak{F}$ respectively are isomorphic. This problem is evidently one which involves the position of $\mathfrak{F}$ in $\mathfrak{B}$; it has a solution if and only if $\mathfrak{F}$ has a special location in $\mathfrak{B}$.

3. **Universal Boolean spaces and rings.** In view of the relations between Boolean spaces and Boolean rings, it is of some interest to consider the algebraic significance of the facts concerning the imbedding of locally-bicompact H-spaces in bicompact H-spaces. We shall consider first the specialization of the theorem to the effect that a locally-bicompact H-space can be imbedded in a bicompact H-space by the adjunction of a single point.* We have

THEOREM 8. *The non-bicompact Boolean spaces are characterized topologically as the non-closed open subsets of bicompact Boolean spaces; in particular, every non-bicompact Boolean space can be converted by the adjunction of a single non-isolated point (in an essentially unique way) into a bicompact Boolean space. Accordingly, the Boolean rings without unit are characterized algebraically as the non-principal ideals in Boolean rings with unit; in particular, every Boolean ring without unit can be imbedded as a non-principal prime ideal in a Boolean ring with unit (in an essentially unique way).*

We have already seen in Theorem 3 that every open subset of a Boolean space is a Boolean space. In a bicompact space a subset is closed if and only if it is bicompact. Hence we see that every non-closed open subset of a bicompact Boolean space is a non-bicompact Boolean space. The converse proposition, that every non-bicompact Boolean space is topologically equivalent to such a subset of some bicompact Boolean space, is a consequence of the assertion concerning the possibility of adjoining a point so as to make the resulting space bicompact. We shall not give a direct topological proof of this assertion, although one could easily be given on the basis of the reference made above to AH. Instead, we deduce it from algebraic considerations. In R Theorems 1 and 37 we have already established the final algebraic statement of the present theorem: every Boolean ring without unit can be imbedded as a non-principal prime ideal in a Boolean ring with unit, which is uniquely determined if isomorphic systems be regarded as identical. Interpreting this fact topologically in accordance with Theorems 4 and 6, we see that every non-bicompact Boolean space is topologically equivalent to an open set obtained from an appropriate bicompact Boolean space by the removal of a single non-isolated point; and we see also that the imbedding space is uniquely determined if equivalent spaces be regarded as identical.

* AH, p. 93, Satz XIV.

Thus, having established all statements of the present theorem except the general algebraic one, we are in a position to prove the latter by a translation of topological facts into algebraic terms. We omit the details.

We pass now to the consideration of a fundamental theorem of Tychonoff,* to the effect that every completely regular space can be imbedded in a certain universal bicompact H-space of the same character. This result suggests a corresponding specialization to the case of Boolean spaces, together with its algebraic interpretation. In formulating such a special theorem, we may disregard the trivial case of finite Boolean spaces: each of them consists of a finite number of discrete points. We therefore state our results in terms of Boolean spaces of infinite character. We have first the following theorem.

THEOREM 9. *Let* $\mathfrak{c}$ *be an arbitrary infinite cardinal number; let* A *be an arbitrary class of cardinal number* $\mathfrak{c}$, *for example, the class of all ordinal numbers preceding some suitable (even the first suitable) ordinal number* ω; *let* $\mathfrak{B}_\mathfrak{c}$ *be the class of all characteristic functions* $\mathfrak{s}=\mathfrak{s}(\alpha)$ *defined over* A *(for each* α *in* A, *either* $\mathfrak{s}(\alpha)=0$ *or* $\mathfrak{s}(\alpha)=1$); *and let* $A_\mathfrak{c}$ *be the class of all sets in* $\mathfrak{B}_\mathfrak{c}$ *generated from the special sets* $\mathfrak{U}_\alpha$, $\mathfrak{U}_\alpha'$, *where* $\mathfrak{U}_\alpha$ *contains all* $\mathfrak{s}$ *for which* $\mathfrak{s}(\alpha)=0$, *by the formation of finite unions and intersections. By the assignment of each non-void set belonging to* $A_\mathfrak{c}$ *as a neighborhood of every one of its points,* $\mathfrak{B}_\mathfrak{c}$ *becomes a bicompact Boolean space of character* $\mathfrak{c}$. *The system* $A_\mathfrak{c}$ *is a Boolean ring with the set* $\mathfrak{B}_\mathfrak{c}$ *as its unit, with* $\mathfrak{c}$ *as its cardinal number, and with* $\mathfrak{B}_\mathfrak{c}$ *as one of its representative Boolean spaces.*

In view of Theorem 1, we can establish the present theorem by showing that $A_\mathfrak{c}$ is a perfect reduced algebra of classes in $\mathfrak{B}_\mathfrak{c}$. From R Theorem 14, we see that $A_\mathfrak{c}$ is a Boolean ring with classes as elements and the particular class $\mathfrak{B}_\mathfrak{c}$ as its unit; and, since the cardinal number of the class of all sets $\mathfrak{U}_\alpha$ is precisely $\mathfrak{c}$, we can easily calculate the cardinal number of $A_\mathfrak{c}$ as equal to $\mathfrak{c}$. This algebra of classes is reduced in the sense of R Definition 10: for it contains the class $\mathfrak{B}_\mathfrak{c}$; and an arbitrary point $\mathfrak{s}_0=\mathfrak{s}_0(\alpha)$ in $\mathfrak{B}_\mathfrak{c}$ is obviously the sole point common to the sets $\mathfrak{U}_\alpha$ where α is such that $\mathfrak{s}_0(\alpha)=0$ and the sets $\mathfrak{U}_\alpha'$ where α is such that $\mathfrak{s}_0(\alpha)=1$. To show that $A_\mathfrak{c}$ is a perfect algebra of classes, we begin by considering an arbitrary prime ideal $\mathfrak{p}$ in $A_\mathfrak{c}$ and defining a corresponding point $\mathfrak{s}_\mathfrak{p}$ in $\mathfrak{B}_\mathfrak{c}$ by setting $\mathfrak{s}_\mathfrak{p}(\alpha)=0$ or $\mathfrak{s}_\mathfrak{p}(\alpha)=1$ according as $\mathfrak{U}_\alpha'\epsilon\mathfrak{p}$ or $\mathfrak{U}_\alpha\epsilon\mathfrak{p}$. We can then prove that no set $\mathfrak{U}$ belonging to $\mathfrak{p}$ contains the point $\mathfrak{s}_\mathfrak{p}$. If $\mathfrak{U}$ is in $\mathfrak{p}$, its complement $\mathfrak{U}'$ is not a member of $\mathfrak{p}$; since $\mathfrak{U}'$ can be represented as a finite union of terms $\mathfrak{U}_{\alpha_1}\cdots\mathfrak{U}_{\alpha_m}\mathfrak{U}_{\beta_1}'\cdots\mathfrak{U}_{\beta_n}'$, where $\alpha_1,\cdots,\alpha_m$, $\beta_1,\cdots,\beta_n$ are in A and either the α's or the β's may be absent, at least one

* Tychonoff, Mathematische Annalen, vol. 102 (1930), pp. 544–561.

of these terms, and indeed every factor of that term, must be a non-member of $\mathfrak{p}$ by virtue of R Theorem 34. Thus the relation $\mathfrak{U}_{\alpha_1}\cdots\mathfrak{U}_{\alpha_m}\mathfrak{U}_{\beta_1}'\cdots\mathfrak{U}_{\beta_n}'\subset\mathfrak{U}'$ can be satisfied by appropriate choice of $\mathfrak{U}_{\alpha_1}',\cdots,\mathfrak{U}_{\alpha_m}',\mathfrak{U}_{\beta_1},\cdots,\mathfrak{U}_{\beta_n}$ in $\mathfrak{p}$. The equations $\mathfrak{s}_{\mathfrak{p}}(\alpha_1)=\cdots=\mathfrak{s}_{\mathfrak{p}}(\alpha_m)=0$, $\mathfrak{s}_{\mathfrak{p}}(\beta_1)=\cdots=\mathfrak{s}_{\mathfrak{p}}(\beta_n)=1$ now show that $\mathfrak{s}_{\mathfrak{p}}$ is contained in $\mathfrak{U}_{\alpha_1}\cdots\mathfrak{U}_{\alpha_m}\mathfrak{U}_{\beta_1}'\cdots\mathfrak{U}_{\beta_n}'$ and in $\mathfrak{U}'$ but not in $\mathfrak{U}$. From R Theorems 59 and 66 it follows therefore that $A_{\mathfrak{c}}$ is a perfect reduced algebra of classes, and is indeed a perfect representation of itself. Theorem 1 shows at once that $\mathfrak{B}_{\mathfrak{c}}$ is a bicompact Boolean space of character $\mathfrak{c}$ which is a representative of $A_{\mathfrak{c}}$ in accordance with Definitions 1 and 2.

THEOREM 10. *Every Boolean space of character not exceeding* $\mathfrak{c}$ *is topologically equivalent to a subspace of the space* $\mathfrak{B}_{\mathfrak{c}}$ *of Theorem* 9; *and every Boolean ring with unit which has cardinal number not exceeding* $\mathfrak{c}$ *is a homomorph of the ring* $A_{\mathfrak{c}}$.

We may obviously restrict attention to a Boolean ring A and its particular representative Boolean space $\mathfrak{E}=\mathfrak{E}(A)$. We may, if we wish, include the case of finite rings. For infinite rings, the cardinal number of A is equal to the character of $\mathfrak{E}$. Hence the assumption that either the cardinal number of A or the character of $\mathfrak{E}(A)$ does not exceed $\mathfrak{c}$ permits us to choose in the class A of Theorem 9 a subset Γ with the same cardinal number as A, and to set up a biunivocal correspondence between A and Γ, designating the element of A associated with the element α in Γ by a_α. In case the cardinal number of A is equal to $\mathfrak{c}$ we may take $\Gamma=\mathrm{A}$ or $\Gamma\neq\mathrm{A}$ as we wish; in all other cases we must take $\Gamma\neq\mathrm{A}$. If now $\mathfrak{p}$ is any point in $\mathfrak{E}$, that is, if $\mathfrak{p}$ is any prime ideal in A, we define a corresponding point $\mathfrak{s}_{\mathfrak{p}}$ in $\mathfrak{B}_{\mathfrak{c}}$ by putting $\mathfrak{s}_{\mathfrak{p}}(\alpha)=0$ if $\alpha\epsilon\mathrm{A}-\Gamma$, $\mathfrak{s}_{\mathfrak{p}}(\alpha)=0$ if $\alpha\epsilon\Gamma$ and $a_\alpha\notin\mathfrak{p}$, and $\mathfrak{s}_{\mathfrak{p}}(\alpha)=1$ if $\alpha\epsilon\Gamma$ and $a_\alpha\epsilon\mathfrak{p}$. We recall that $\mathfrak{p}\epsilon\mathfrak{E}(a_\alpha)$ if and only if $a_\alpha\notin\mathfrak{p}$. The class of all points $\mathfrak{s}_{\mathfrak{p}}$ thus obtained will be denoted by $\mathfrak{B}_{\mathfrak{c}}(A)$. It is evident that the correspondence between $\mathfrak{E}$ and $\mathfrak{B}_{\mathfrak{c}}(A)$ is biunivocal. In order that a subset of $\mathfrak{B}_{\mathfrak{c}}(A)$ be the image of $\mathfrak{E}(a_\beta)$, it is necessary and sufficient that its points $\mathfrak{s}_{\mathfrak{p}}$ have indices $\mathfrak{p}$ such that $\mathfrak{p}\epsilon\mathfrak{E}(a_\beta)$ or, equivalently, such that $\mathfrak{s}_{\mathfrak{p}}(\beta)=0$. Thus the images of the sets $\mathfrak{E}(a_\beta)$ in $\mathfrak{E}$ are precisely the sets $\mathfrak{U}_\beta\mathfrak{B}_{\mathfrak{c}}(A)$, where $\beta\epsilon\Gamma$. It is evident that $\alpha\epsilon\mathrm{A}-\Gamma$ implies $\mathfrak{U}_\alpha\supset\mathfrak{B}_{\mathfrak{c}}(A)$. Now the intersections of the sets in $A_{\mathfrak{c}}$ with $\mathfrak{B}_{\mathfrak{c}}(A)$ constitute a basis in $\mathfrak{B}_{\mathfrak{c}}(A)$; indeed, by the definition of the class $A_{\mathfrak{c}}$, we may restrict attention to the sets $\mathfrak{U}_{\alpha_1}\cdots\mathfrak{U}_{\alpha_m}\mathfrak{U}_{\beta_1}'\cdots\mathfrak{U}_{\beta_n}'$ which have points in common with $\mathfrak{B}_{\mathfrak{c}}(A)$, it being understood that either the α's or the β's may be absent. Under the indicated restriction we may suppose, by virtue of the previous observations, that the indices $\alpha_1,\cdots,\alpha_m,\beta_1,\cdots,\beta_n$ belong to Γ. The sets $\mathfrak{U}_{\alpha_1}\cdots\mathfrak{U}_{\alpha_m}\mathfrak{U}_{\beta_1}'\cdots\mathfrak{U}_{\beta_n}'\mathfrak{B}_{\mathfrak{c}}(A)$ are then precisely the images of the sets $\mathfrak{E}(a_{\alpha_1})\cdots\mathfrak{E}(a_{\alpha_m})\mathfrak{E}'(a_{\beta_1})\cdots\mathfrak{E}'(a_{\beta_n})=\mathfrak{E}(a_\gamma)\mathfrak{E}'(a_\delta)$ in $\mathfrak{E}$,

where $a_\gamma = a_{\alpha_1} \cdots a_{\alpha_m}$ and $a_\delta = a_{\beta_1} \vee \cdots \vee a_{\beta_n}$. Since the latter sets obviously constitute a basis in $\mathfrak{E}$, we see that the correspondence set up between $\mathfrak{E}$ and $\mathfrak{B}_\mathfrak{c}(A)$ is a topological equivalence. When A has a unit, $\mathfrak{E}$ and $\mathfrak{B}_\mathfrak{c}(A)$ are bicompact; and $\mathfrak{B}_\mathfrak{c}(A)$ is therefore closed in $\mathfrak{B}_\mathfrak{c}$. By Theorem 4 it follows that A is a homomorph of $A_\mathfrak{c}$.

The theorem just proved is the desired specialization of the theorem of Tychonoff. From it we can obtain further useful facts.

THEOREM 11. *The Boolean spaces $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}\}$ obtained from $\mathfrak{B}_\mathfrak{c}$ by the suppression of a single point $\mathfrak{s}$ are topologically equivalent and all non-bicompact; the prime ideals in the Boolean ring $A_\mathfrak{c}$ are isomorphic Boolean rings without unit. If any common isomorph of the prime ideals in $A_\mathfrak{c}$ be denoted by $A_\mathfrak{c}^*$, then every Boolean ring of cardinal number not exceeding $\mathfrak{c}$ is a homomorph of $A_\mathfrak{c}^*$.*

The relations $\mathfrak{s}' = \mathfrak{s} + \mathfrak{s}_0$ (mod 2), $\mathfrak{s}_0 = \mathfrak{s}_2 - \mathfrak{s}_1$ (mod 2) obviously set up a biunivocal correspondence of $\mathfrak{B}_\mathfrak{c}$ with itself, carrying $\mathfrak{s}_1$ into $\mathfrak{s}' = \mathfrak{s}_2$. It is evident that the image of the set $\mathfrak{U}_\alpha$ under this correspondence is the set specified by $\mathfrak{s}'(\alpha) = \mathfrak{s}_0(\alpha)$ and hence that the image is either $\mathfrak{U}_\alpha$ or $\mathfrak{U}_\alpha'$ according as $\mathfrak{s}_0(\alpha) = 0$ or $\mathfrak{s}_0(\alpha) = 1$. It follows readily that the correspondence transforms $A_\mathfrak{c}$ into itself and is therefore a topological transformation of $\mathfrak{B}_\mathfrak{c}$ into itself. In consequence, the spaces $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}_1\}$ and $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}_2\}$ are topologically equivalent. It is clear that no point of $\mathfrak{B}_\mathfrak{c}$ is isolated: this can be verified directly; or can be proved by noting that, if one point were isolated, then every point would be isolated in accordance with the fact that the topological group of $\mathfrak{B}_\mathfrak{c}$ is transitive, but in contradiction with the fact that $\mathfrak{B}_\mathfrak{c}$ is bicompact. We see therefore that every space $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}\}$ is a non-bicompact Boolean space. By Theorems 4 and 6, these results imply that the prime ideals in $A_\mathfrak{c}$ have the asserted properties. If we recall the relations between prime ideals and atomic elements described in R Theorem 38, we see in particular that $A_\mathfrak{c}$ contains no atomic element and no normal prime ideal.

It is now easily verified that every Boolean ring A of cardinal number not exceeding $\mathfrak{c}$ is a homomorph of $A_\mathfrak{c}^*$. We first adjoin a unit to A, if necessary, obtaining a Boolean ring A^*. This adjunction can be effected by means of R Theorem 1. It is easily seen by reference to the construction given there that the cardinal number of A^* likewise does not exceed $\mathfrak{c}$. We now proceed, as in the proof of Theorem 10, to construct a representative $\mathfrak{B}_\mathfrak{c}(A^*)$ of A^*, taking Γ different from A so that $\mathfrak{B}_\mathfrak{c}(A^*)$ is a proper subset of $\mathfrak{B}_\mathfrak{c}$. When A has a unit, we have $A = A^*$ and know that $\mathfrak{B}_\mathfrak{c}(A^*)$ is a closed subset of $\mathfrak{B}_\mathfrak{c}$. Thus, if we choose $\mathfrak{s}$ as a point in $\mathfrak{B}_\mathfrak{c}'(A^*)$, the space $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}\}$ contains $\mathfrak{B}_\mathfrak{c}(A^*)$ as a closed subset. In this case, $A = A^*$ is a homomorph of $A_\mathfrak{c}^*$ in accordance with Theorem 4. When A has no unit, A is a prime ideal in A^* by R The-

orem 37. It follows that a representative $\mathfrak{B}_\mathfrak{c}(A)$ of A can be found by removing an appropriate point $\mathfrak{s}$ from the space $\mathfrak{B}_\mathfrak{c}(A^*)$. Since we have $\mathfrak{B}_\mathfrak{c}(A)=\mathfrak{B}_\mathfrak{c}(A^*)(\mathfrak{B}_\mathfrak{c}-\{\mathfrak{s}\})$, the set $\mathfrak{B}_\mathfrak{c}(A)$ is closed in the space $\mathfrak{B}_\mathfrak{c}-\{\mathfrak{s}\}$. Hence in this case also, A is a homomorph of $A_\mathfrak{c}^*$.

Just as the Boolean space $\mathfrak{B}_\mathfrak{c}$ is a universal Boolean space in the sense that in it can be imbedded every Boolean space of sufficiently small character, so the Boolean rings $A_\mathfrak{c}$ and $A_\mathfrak{c}^*$ are universal Boolean rings in the sense that their homomorphs exhaust all Boolean rings of cardinal number not exceeding $\mathfrak{c}$, due regard being paid in the case of $A_\mathfrak{c}$ to the condition that the homomorphs of a ring with unit all possess units. If we recall that the homomorphs of a Boolean ring can be obtained through the replacement of the fundamental equality by the various congruences in the ring, we are led to formulate the following algebraic characterization of the universal rings $A_\mathfrak{c}$ and $A_\mathfrak{c}^*$:

THEOREM 12. *The Boolean rings $A_\mathfrak{c}$ and $A_\mathfrak{c}^*$ are isomorphic respectively to the free Boolean rings with and without unit generated by $\mathfrak{c}$ elements.*

The free Boolean ring without unit generated by $\mathfrak{c}$ elements is obtained by forming all abstract or symbolic polynomials in these elements and then introducing the weakest possible relation of equality consistent with the postulates for a Boolean ring without unit. In accordance with this procedure the symbolic polynomials $a+a$, where a is a generating element, must all be equated; and we can denote some one of them by 0. The elements of the free ring which are equal to 0 then have in common the properties of a zero-element. If we make use of the postulates for Boolean rings, we can describe the introduction of the equality between symbolic polynomials in a more precise way: the desired equality is determined uniquely by the property that a symbolic polynomial is equated to 0 if and only if the algebraic laws expressed or implied in the fundamental postulates for a Boolean ring without unit reduce it formally to 0. We thus see that the free ring can be characterized as follows: it is a Boolean ring without unit; it contains a subset of cardinal number $\mathfrak{c}$ which generates the entire ring; and the elements of this subset satisfy no algebraic relation which is not a Boolean identity.

The free Boolean ring with unit generated by $\mathfrak{c}$ elements is obtained by forming the free ring without unit generated by the given elements and then adjoining a unit by the construction of R Theorem 1. Thus this free ring can be characterized as follows: it is a Boolean ring with unit; it contains a subset of cardinal number $\mathfrak{c}$ which, together with the unit, generates the entire ring; and the elements of this subset satisfy no algebraic relation which is not a Boolean identity.

In order to apply the given characterizations of free rings, we need a criterion to determine whether or not the elements of a subclass $\mathfrak{a}$ of a Boolean ring A are free from non-identical algebraic relations. Such a criterion can be established as follows: in order that the elements of $\mathfrak{a}$ shall satisfy no algebraic relation which is not a Boolean identity, it is necessary and sufficient that no relation $a_1 \cdots a_m < b_1 \vee \cdots \vee b_n$, $m \geqq 1$, $n \geqq 1$ shall hold between distinct elements $a_1, \cdots, a_m, b_1, \cdots, b_n$ in $\mathfrak{a}$. We first suppose that A has a unit. Then any non-identical algebraic relation connecting distinct elements of $\mathfrak{a}$ can be reduced to the form $d_1 \vee \cdots \vee d_p = 0$, where each element d is expressed in terms of distinct elements of $\mathfrak{a}$ by a relation $d = a_1 \cdots a_m b_1' \cdots b_n'$. The given relation therefore implies $d_1 = \cdots = d_p = 0$ and hence $a_1 \cdots a_m b_1' \cdots b_n' = 0$. We may obviously suppose that $m \geqq 1$, $n \geqq 1$: for, if $b_1' \cdots b_n' = 0$ or $a_1 \cdots a_m = 0$, the missing elements $a_1, \cdots, a_m$ or $b_1', \cdots, b_n'$ can be inserted arbitrarily. Thus we have shown that any non-identical algebraic relation connecting distinct elements of $\mathfrak{a}$ implies a relation $a_1 \cdots a_m b_1' \cdots b_n' = 0$ or the equivalent relation $a_1 \cdots a_m < b_1 \vee \cdots \vee b_n$, $m \geqq 1$, $n \geqq 1$, connecting distinct elements of $\mathfrak{a}$. In case A has no unit, we adjoin one by the construction of R Theorem 1 and repeat the argument just given. It is thus evident in all cases that $\mathfrak{a}$ is free from non-identical algebraic relations if and only if no relation $a_1 \cdots a_m < b_1 \vee \cdots \vee b_n$ holds between distinct elements of $\mathfrak{a}$.

The Boolean ring $A_\mathfrak{c}$ is evidently generated by the class of all sets $\mathfrak{U}_\alpha$ together with the unit $\mathfrak{B}_\mathfrak{c}$. This class of sets obviously has cardinal number $\mathfrak{c}$. We now observe that the relation $\mathfrak{U}_{\alpha_1} \cdots \mathfrak{U}_{\alpha_m} \subset \mathfrak{U}_{\beta_1} \cup \cdots \cup \mathfrak{U}_{\beta_n}$, $m \geqq 1$, $n \geqq 1$, means that every characteristic function $\mathfrak{s} = \mathfrak{s}(\alpha)$ with the property $\mathfrak{s}(\alpha_1) = \cdots = \mathfrak{s}(\alpha_m) = 0$ has also at least one of the properties $\mathfrak{s}(\beta_1) = 0, \cdots, \mathfrak{s}(\beta_n) = 0$. Obviously, then, no such relation can hold when the indices $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n$ are distinct. It follows that $A_\mathfrak{c}$ is isomorphic to the free Boolean ring with unit generated by $\mathfrak{c}$ elements.

The Boolean ring $A_\mathfrak{c}^*$ may be regarded as that prime ideal in $A_\mathfrak{c}$ which contains every set $\mathfrak{U}_\alpha$. In fact, we can determine a prime ideal in $A_\mathfrak{c}$ by considering those sets in $A_\mathfrak{c}$ which are contained in $\mathfrak{B}_\mathfrak{c} - \{\mathfrak{s}\}$, where $\mathfrak{s}(\alpha) = 1$ for every α. It is evident that this prime ideal contains every set $\mathfrak{U}_\alpha$. Since $A_\mathfrak{c}^*$ is isomorphic to any prime ideal in $A_\mathfrak{c}$, we may identify it with the one just determined. It is then evident that the ring $A_\mathfrak{c}^*$ is generated by the class of all sets $\mathfrak{U}_\alpha$. As in the preceding paragraph, it follows that $A_\mathfrak{c}^*$ is isomorphic to the free Boolean ring without unit generated by $\mathfrak{c}$ elements.

In conclusion, it is of some interest to remark upon the duality between the free Boolean ring with unit generated by $\mathfrak{c}$ elements and the Boolean ring of all subclasses of a fixed class of cardinal number $\mathfrak{c}$ which emerges

from the previous discussion. Since the characteristic functions $\mathfrak{s}(\alpha)$ constitute under multiplication and addition (mod 2) a Boolean ring isomorphic to the ring of all subclasses of A, our construction of the free ring $A_{\mathfrak{c}}$ in terms of these characteristic functions may be regarded as exhibiting a dual relationship between the rings in question. This duality is in fact a special instance of that which holds between discrete abelian groups and the subgroups of toroidal groups:* for Boolean rings are abelian groups under addition and can be studied from that point of view. We shall not pursue this aspect of the subject further.

We shall close the present section with an examination of the special case $\mathfrak{c}=\aleph_0$. We obtain a result which explains in some measure the frequent occurrence of the Cantor discontinuum in various branches of mathematics.

THEOREM 13. *The Boolean space $\mathfrak{B}_{\mathfrak{c}}$, $\mathfrak{c}=\aleph_0$, is topologically equivalent to the space $\mathfrak{D}$ known as the Cantor discontinuum.*

$\mathfrak{D}$ may be described as the set of all real numbers x given by developments $x=2\sum_{\alpha=1}^{\infty}3^{-\alpha}s(\alpha)$, where $s(k)=0$ or $s(k)=1$, with the usual metric topology. An important property of $\mathfrak{D}$ is the following: a sequence $\{x_n\}$ in $\mathfrak{D}$ converges if and only if $\{s_n(k)\}$ converges for each k; and the limit of $\{x_n\}$ is obtained as the element x with development given by $s(k)=\lim_{n\to\infty} s_n(k)$, $k=1, 2, 3, \cdots$. To prove this assertion, we consider the inequality

$$\left| 2\sum_{\alpha=1}^{N} 3^{-\alpha}(s_m(\alpha) - s_n(\alpha)) \right| \leqq 2 \sum_{\alpha=N+1}^{\infty} 3^{-\alpha} | s_m(\alpha) - s_n(\alpha) | + | x_m - x_n |$$
$$\leqq 3^{-N} + | x_m - x_n | .$$

If $\{x_n\}$ is convergent and N is the first integer such that $\{s_n(N)\}$ is not known to converge, this inequality becomes

$$2\cdot 3^{-N} | s_m(N) - s_n(N) | \leqq 3^{-N} + | x_m - x_n |$$

for all sufficiently large indices m and n. If we restrict m and n to be so great that $|x_m - x_n| < 3^{-N}$, the inequality reduces to $|s_m(N)-s_n(N)|<1$. Hence we must have $s_m(N)=s_n(N)$ for all sufficiently large indices m and n. Since this result establishes the convergence of $\{s_n(N)\}$, we can apply the principle of mathematical induction to conclude that $\{s_n(k)\}$ converges for every k. If we know that $s(k)=\lim_{n\to\infty} s_n(k)$ exists for every k, we see readily that the number $x=2\sum_{\alpha=1}^{\infty}3^{-\alpha}s(\alpha)$ has the property

$$| x - x_n | \leqq 2\sum_{\alpha=1}^{\infty} 3^{-\alpha} | s(\alpha) - s_n(\alpha) | \leqq 2\sum_{\alpha=1}^{N} 3^{-\alpha} | s(\alpha) - s_n(\alpha) | + 3^{-N} < \epsilon$$

* Alexander and Zippin, Annals of Mathematics, (2), vol. 36 (1935), pp. 71–85.

for $N > -\log \epsilon/\log 3$, $n > M = M(N)$. Combining these results in the obvious way, we obtain the proposition stated above. There are several useful consequences of this proposition. First, it shows that if x and x' have developments given by $s(k)$ and $s'(k)$ respectively, then $x = x'$ implies $s(k) = s'(k)$ for every k: indeed, the sequence $x, x', x, x', \cdots$ is then convergent so that the sequence $s(k), s'(k), s(k), s'(k), \cdots$ must converge likewise, for every k. Secondly, it shows that the set of points x obtained by fixing the values of $s(k)$ for $k = 1, \cdots, N$ and leaving the values of $s(k)$ for $k \geqq N+1$ arbitrary is an open subset of $\mathfrak{D}$: for the complement of this set in $\mathfrak{D}$ has the property that any convergent subsequence has a limit for which one of the development-coefficients $s(k)$, $k = 1, \cdots, N$, has a value different from the one prescribed; in other words, the complement of the given set is closed in $\mathfrak{D}$. Thirdly, it shows that $\mathfrak{D}$ is bicompact: for any convergent sequence in $\mathfrak{D}$ has a limit in $\mathfrak{D}$; and $\mathfrak{D}$ is thus a closed subset of the bicompact space consisting of the real numbers x, where $0 \leqq x \leqq 1$, with the usual topology. The correspondence given by $x \rightarrow \mathfrak{s} = s(k)$ is a biunivocal correspondence between $\mathfrak{D}$ and $\mathfrak{B}_\mathfrak{c}$, $\mathfrak{c} = \aleph_0$, by virtue of our first remark. The sets $\mathfrak{U}_{\alpha_1} \cdots \mathfrak{U}_{\alpha_m} \mathfrak{U}_{\beta_1}' \cdots \mathfrak{U}_{\beta_n}'$ where $m + n = N$ and $(\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n)$ is a permutation of $(1, \cdots, N)$ obviously constitute a basis in $\mathfrak{B}_\mathfrak{c}$. Their antecedents in $\mathfrak{D}$ are precisely the sets obtained by fixing the first N coefficients in the development of x in each of the 2^N possible ways; and, since these antecedents are open in $\mathfrak{D}$ by virtue of our second remark, the correspondence from $\mathfrak{D}$ to $\mathfrak{B}_\mathfrak{c}$, $\mathfrak{c} = \aleph_0$ is continuous. Finally, since $\mathfrak{D}$ and $\mathfrak{B}_\mathfrak{c}$, $\mathfrak{c} = \aleph_0$, are both bicompact H-spaces, the correspondence is necessarily bicontinuous; and $\mathfrak{D}$ and $\mathfrak{B}_\mathfrak{c}$, $\mathfrak{c} = \aleph_0$, are topologically equivalent.*

Chapter II. Maps in Boolean spaces

1. **The general theory of maps.** The general theory of maps deals with the problem of representing the points and properties of a given topological space by subsets and their properties in a second topological space. While we shall be concerned primarily with the study of such representations in Boolean spaces, we shall begin with an elementary discussion of quite general maps, introducing the terminology appropriate to the later theory.

The point of departure for the entire theory is the following result:

Theorem 14. *If $\mathfrak{X}$ is any non-void family of distinct, but not necessarily disjoint, non-void closed sets $\mathfrak{x}$ in a T_1-space $\mathfrak{S}$, then a topology† may be imposed on $\mathfrak{X}$ by the introduction of the following system of neighborhoods: each*

* AH, p. 95, Satz III.

† Compare with AH, p. 66, where the sets $\mathfrak{x}$ are assumed to be disjoint and the topology is termed "weak" to distinguish it from a different topology introduced earlier on pp. 61–62.

non-void subfamily of $\mathfrak{X}$ *which is characterized as the class of all* $\mathfrak{x}$ *contained in a fixed open subset* $\mathfrak{G}$ *of* $\mathfrak{S}$ *is assigned as a neighborhood of every one of its elements* $\mathfrak{x}$. *Under this topology,* $\mathfrak{X}$ *is a* T_0*-space. In order that* $\mathfrak{X}$ *be a* T_1*-space, it is necessary and sufficient that no member of* $\mathfrak{X}$ *contain another as a proper subset.*

It is evident that the system of neighborhoods described in the theorem has the following properties: every "point" $\mathfrak{x}$ belongs to each of its neighborhoods; the intersection of any two neighborhoods of an arbitrary point is also a neighborhood of that point; if an arbitrary neighborhood contains a given point, it is a neighborhood of that point; and every point has at least one neighborhood. Furthermore, we can show that, if $\mathfrak{x}_1$ and $\mathfrak{x}_2$ are distinct points, then at least one of them has a neighborhood which does not contain the other. Since $\mathfrak{x}_1$ and $\mathfrak{x}_2$ are distinct as subsets of $\mathfrak{S}$, we may suppose our notation so chosen that $\mathfrak{x}_2$ contains a point $\mathfrak{s}_2$ of $\mathfrak{S}$ which does not belong to $\mathfrak{x}_1$. Since $\mathfrak{S}$ is a T_1-space, every $\mathfrak{s}_1$ in $\mathfrak{x}_1$ is contained in an open set $\mathfrak{G}(\mathfrak{s}_1)$ which does not contain $\mathfrak{s}_2$; and the union of all the open sets $\mathfrak{G}(\mathfrak{s}_1)$, $\mathfrak{s}_1\epsilon\mathfrak{x}_1$, is an open set $\mathfrak{G}$ which contains $\mathfrak{x}_1$ but does not contain either $\mathfrak{s}_2$ or the set $\mathfrak{x}_2$. It is then evident that the family specified by $\mathfrak{x}\subset\mathfrak{G}$ is a neighborhood of $\mathfrak{x}_1$ which does not contain $\mathfrak{x}_2$. From this argument we see that, if neither of the distinct sets $\mathfrak{x}_1$ and $\mathfrak{x}_2$ contains the other, then each has, as a point in $\mathfrak{X}$, a neighborhood which does not contain the other; and also that, if one contains the other, as a proper subset, then every one of its neighborhoods contains the second. These facts concerning the system of neighborhoods identify $\mathfrak{X}$ as a T_0-space, and show that $\mathfrak{X}$ is a T_1-space if and only if no set $\mathfrak{x}$ contains another as a proper subset.*

It is convenient, in view of further developments, to introduce a certain amount of descriptive terminology relating to the family $\mathfrak{X}$. We therefore give the following definitions:

DEFINITION 3. *If the union* $\mathfrak{S}(\mathfrak{X})$ *of the sets belonging to* $\mathfrak{X}$ *contains a subset* $\mathfrak{A}$ *of* $\mathfrak{S}$, *then* $\mathfrak{X}$ *is said to cover* $\mathfrak{A}$.

DEFINITION 4. *If every non-void open set* $\mathfrak{G}$ *in* $\mathfrak{S}$ *contains some member of* $\mathfrak{X}$ *as a subset, then* $\mathfrak{X}$ *is said to be densely distributed in* $\mathfrak{S}$.

DEFINITION 5. *If, whenever* $\mathfrak{G}$ *is an open set in* $\mathfrak{S}$ *and* $\mathfrak{x}_0$ *is a member of* $\mathfrak{X}$, *the relation* $\mathfrak{G}\supset\mathfrak{x}_0$ *implies the existence of an open set* $\mathfrak{G}_0$ *such that* $\mathfrak{G}_0\supset\mathfrak{x}_0$ *and* $\mathfrak{x}\subset\mathfrak{G}$ *whenever* $\mathfrak{x}\mathfrak{G}_0\neq 0$, *then* $\mathfrak{X}$ *is called a continuous family.*

DEFINITION 6. *If* $\mathfrak{c}$ *is the least cardinal number such that* $\mathfrak{c}+1$ *is exceeded by the cardinal number of no set* $\mathfrak{x}$ *in* $\mathfrak{X}$, *then* $\mathfrak{c}$ *is called the index of the family* $\mathfrak{X}$.

* AH, pp. 58–59.

If some set $\mathfrak{X}$ *in* X *has cardinal number* $\mathfrak{c}+1$, *the index is said to be attained.*

We proceed now to give the principal definitions dealing with the general concept of a map.

DEFINITION 7. *If* $\mathfrak{R}$ *is a topological space equivalent to the* T_0*-space* X *of Theorem* 14, *then the relation between* $\mathfrak{R}$, $\mathfrak{S}$, *and* X *is said to be a map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *of* $\mathfrak{R}$ *in* $\mathfrak{S}$ *defined by the family* X. *A map defined by a subfamily of* X *is called a submap of* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$. *A map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *is said to be a covering map or a map of* $\mathfrak{R}$ *on* $\mathfrak{S}$ *if* X *covers* $\mathfrak{S}$. *A map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *is said to be continuous if* X *is a continuous family. The index of the family* X *is called the index of the map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$. *A map is said to be a Boolean map if the space* $\mathfrak{S}$ *is a bicompact Boolean space.*

It is convenient to state explicitly the following elementary theorems.

THEOREM 15. *If* $\mathfrak{R}_1$, $\mathfrak{S}_1$ *are topologically equivalent to* $\mathfrak{R}_2$, $\mathfrak{S}_2$ *respectively, then the existence of a map* $m(\mathfrak{R}_1, \mathfrak{S}_1, \mathsf{X}_1)$ *implies and is implied by the existence of a map* $m(\mathfrak{R}_2, \mathfrak{S}_2, \mathsf{X}_2)$ *whenever the equivalence between* $\mathfrak{S}_1$ *and* $\mathfrak{S}_2$ *carries the families* X_1 *and* X_2 *into one another.*

THEOREM 16. *If* $\mathfrak{S}$ *is a closed subset of a* T_1*-space* $\mathfrak{T}$ *and* X *is a family of subsets of* $\mathfrak{S}$, *then the members of* X *are closed relative to* $\mathfrak{S}$ *if and only if they are closed relative to* $\mathfrak{T}$; *and the existence of a map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *defined by the family* X *in* $\mathfrak{S}$ *implies and is implied by the existence of a map* $m(\mathfrak{R}, \mathfrak{T}, \mathsf{X})$ *defined by the family* X *in* $\mathfrak{T}$. *If* $\mathfrak{S}$ *is any subset of a* T_1*-space* $\mathfrak{T}$ *and* X *is a family of subsets of* $\mathfrak{S}$ *closed relative to* $\mathfrak{T}$, *then the members of* X *are closed relative to* $\mathfrak{S}$; *and the existence of a map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *implies and is implied by the existence of a map* $m(\mathfrak{R}, \mathfrak{T}, \mathsf{X})$. *In particular, the existence of a map* $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ *implies and is implied by the existence of a map* $m(\mathfrak{R}, \mathfrak{S}(\mathsf{X}), \mathsf{X})$ *or a map* $m(\mathfrak{R}, \mathfrak{S}^-(\mathsf{X}), \mathsf{X})$, *where* X *is a family of distinct non-void closed subsets of* $\mathfrak{S}$.

In accordance with Theorem 15, we introduce the following definition.

DEFINITION 8. *If two maps* $m(\mathfrak{R}_1, \mathfrak{S}_1, \mathsf{X}_1)$ *and* $m(\mathfrak{R}_2, \mathfrak{S}_2, \mathsf{X}_2)$ *are related in the manner described in Theorem* 15, *they are said to be equivalent.*

It is evident that equivalent maps are topologically indistinguishable. In particular, if one of two equivalent maps is continuous or covering, then the other is also; and two equivalent maps have the same index. For purposes of exposition, it is often convenient to recall that, by definition, a map $m(\mathfrak{R}, \mathfrak{S}, \mathsf{X})$ is equivalent to the "identical" map $m(\mathsf{X}, \mathfrak{S}, \mathsf{X})$ in which the points of the space X of Theorem 14 are represented by themselves as closed subsets of $\mathfrak{S}$.

We can now raise three general questions concerning maps: first, we in-

quire what consequences flow from the imposition of various simple conditions upon $\mathfrak{R}$, $\mathfrak{S}$, or $\mathcal{X}$ in the map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$; second, we inquire what reductions can be applied to a map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$ in order to obtain a simplified map $m(\mathfrak{R}, \mathfrak{S}^*, \mathcal{X}^*)$; and, third, we inquire what connections link the theory of maps with the theory of topological images. We shall consider these three problems in succession.

In Theorem 14, we have already seen that in a map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$ or its equivalent $m(\mathcal{X}, \mathfrak{S}, \mathcal{X})$ the space $\mathfrak{R}$ is a T_1-space if and only if the family $\mathcal{X}$ has a certain simple property. Beside this result we may now place the following theorem.

THEOREM 17. *In a map* $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$, *the condition that the members of* $\mathcal{X}$ *be disjoint sets is neither necessary nor sufficient for* $\mathfrak{R}$ *to be an* H-*space; but this condition becomes sufficient in the presence of either of the auxiliary conditions*:

(1) $\mathfrak{S}$ *is a normal space*;

(2) $\mathfrak{S}$ *is an* H-*space and the sets in* $\mathcal{X}$ *are bicompact.*

Let us first consider the sufficiency of the indicated condition when (1) or (2) is valid. If $\mathfrak{S}$ is normal and $\mathfrak{X}_1$ and $\mathfrak{X}_2$ are disjoint members of $\mathcal{X}$, the definition of normality establishes the existence of disjoint open sets $\mathfrak{G}_1$ and $\mathfrak{G}_2$ in $\mathfrak{S}$ such that the closed sets $\mathfrak{X}_1$, $\mathfrak{X}_2$ satisfy the relations $\mathfrak{X}_1 \subset \mathfrak{G}_1$, $\mathfrak{X}_2 \subset \mathfrak{G}_2$. Hence any two such points in $\mathcal{X}$ have disjoint neighborhoods in the neighborhood system of Theorem 14. It follows that our condition is sufficient for $\mathcal{X}$ and $\mathfrak{R}$ to be H-spaces. Under condition (2), we construct analogous open sets $\mathfrak{G}_1$ and $\mathfrak{G}_2$ by the method used to show that every bicompact H-space is normal. If $\mathfrak{X}_1$ and $\mathfrak{X}_2$ are disjoint and bicompact in the H-space $\mathfrak{S}$, there exist disjoint open sets $\mathfrak{G}_1(\mathfrak{s}_1, \mathfrak{s}_2)$, $\mathfrak{G}_2(\mathfrak{s}_1, \mathfrak{s}_2)$ containing respectively an arbitrary point $\mathfrak{s}_1$ in $\mathfrak{X}_1$ and an arbitrary point $\mathfrak{s}_2$ in $\mathfrak{X}_2$; for fixed $\mathfrak{s}_2$ the sets $\mathfrak{G}_1(\mathfrak{s}_1, \mathfrak{s}_2)$ cover $\mathfrak{X}_1$ so that there exist points $\mathfrak{s}_1^{(1)}, \cdots, \mathfrak{s}_1^{(m)}$, dependent upon $\mathfrak{s}_2$, for which $\mathfrak{G}_1(\mathfrak{s}_2) = \mathfrak{G}_1(\mathfrak{s}_1^{(1)}, \mathfrak{s}_2) \cup \cdots \cup \mathfrak{G}_1(\mathfrak{s}_1^{(m)}, \mathfrak{s}_2) \supset \mathfrak{X}_1$; since the open sets $\mathfrak{G}_2(\mathfrak{s}_2) = \mathfrak{G}_2(\mathfrak{s}_1^{(1)}, \mathfrak{s}_2) \cdots \mathfrak{G}_2(\mathfrak{s}_1^{(m)}, \mathfrak{s}_2)$ cover $\mathfrak{X}_2$, there exist points $\mathfrak{s}_2^{(1)}, \cdots, \mathfrak{s}_2^{(n)}$ such that $\mathfrak{G}_2 = \mathfrak{G}_2(\mathfrak{s}_2^{(1)}) \cup \cdots \cup \mathfrak{G}_2(\mathfrak{s}_2^{(n)}) \supset \mathfrak{X}_2$; since the open set $\mathfrak{G}_1(\mathfrak{s}_2^{(k)})$ is disjoint from $\mathfrak{G}_2(\mathfrak{s}_2^{(k)})$ and contains $\mathfrak{X}_1$ for $k = 1, \cdots, n$, the open sets $\mathfrak{G}_1 = \mathfrak{G}_1(\mathfrak{s}_2^{(1)}) \cdots \mathfrak{G}_1(\mathfrak{s}_2^{(n)})$ and $\mathfrak{G}_2$ are disjoint and contain $\mathfrak{X}_1$ and $\mathfrak{X}_2$ respectively. Again it follows that our condition is sufficient for $\mathfrak{X}$ and $\mathfrak{R}$ to be H-spaces.

On the other hand, if $\mathfrak{S}$ is not normal, it is possible for two disjoint closed sets $\mathfrak{X}_1$ and $\mathfrak{X}_2$ to have the property that, whatever the open sets $\mathfrak{G}_1$ and $\mathfrak{G}_2$ containing $\mathfrak{X}_1$ and $\mathfrak{X}_2$ respectively, their intersection $\mathfrak{G}_1\mathfrak{G}_2$ is non-void. In such a situation we cannot conclude in general that the points $\mathfrak{X}_1$ and $\mathfrak{X}_2$ in $\mathcal{X}$ have disjoint neighborhoods. Yet, if $\mathfrak{X}_1$ and $\mathfrak{X}_2$ have common points, they

may still have disjoint neighborhoods, as can be seen from the following example: let $\mathcal{X}$ consist of two closed sets $\mathfrak{X}_1$ and $\mathfrak{X}_2$ which have a common point but neither of which contains the other; then $\mathcal{X}$ is an H-space.

The further examination of the effect of restrictions upon $\mathfrak{S}$ and $\mathcal{X}$ is, in a sense, the chief aim of the following sections. Indeed, we shall see that the assumption that $\mathfrak{S}$ is a bicompact Boolean space has no effect upon the topological nature of the spaces $\mathfrak{R}$ which can be mapped in $\mathfrak{S}$. The relation of the properties of $\mathcal{X}$ to those of $\mathfrak{R}$ is found to be rather complex. For the present, then, we leave the first problem with the remark that every bicompact Boolean space is normal and its closed subsets bicompact, thus rendering Theorem 17 applicable to such spaces.

We pass to our second problem. In Theorem 16 we have already noted that a map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$ can be modified, without the loss of essential information, through the suppression of $\mathfrak{S}'(\mathcal{X})$ or of $\mathfrak{S}^{-\prime}(\mathcal{X})$. It will be observed that the second of these sets is an open set disjoint from all members of $\mathcal{X}$ and that it is the maximal set with such properties. Now it is possible for an open set $\mathfrak{H}$ in $\mathfrak{S}$ to have points in common with some members of $\mathcal{X}$ and yet contain no member of $\mathcal{X}$. It is natural to examine the conditions under which such a set $\mathfrak{H}$ can be suppressed from a given map without disturbing the topological relations involved. We obtain the following result.

THEOREM 18. *Let $\mathfrak{H}$ be an arbitrary open subset of a T_1-space $\mathfrak{S}$; let $\mathcal{X}$ be a family of distinct closed sets $\mathfrak{X}$ in $\mathfrak{S}$, none of which is contained in $\mathfrak{H}$; let $\mathfrak{T}$ be the set $\mathfrak{H}'$ considered as a relative space; and let Υ be the family of the closed sets $\mathfrak{Y}=\mathfrak{H}'\mathfrak{X}$ in $\mathfrak{T}$. Under the topology introduced in Theorem* 14, *the spaces $\mathcal{X}$ and Υ have the property that Υ is a continuous image of $\mathcal{X}$ by virtue of the correspondence $\mathfrak{X}\rightarrow\mathfrak{Y}=\mathfrak{H}'\mathfrak{X}$. In order that $\mathcal{X}$ and Υ be topologically equivalent under this correspondence, it is necessary and sufficient that $\mathfrak{H}$ have the following property: if $\mathfrak{G}$ is any open set in $\mathfrak{S}$ and $\mathfrak{X}_0$ any member of $\mathcal{X}$ contained in $\mathfrak{G}$, then there exists an open set $\mathfrak{G}_0$ such that $\mathfrak{G}_0 \cup \mathfrak{H}$ contains $\mathfrak{X}_0$ and contains no member $\mathfrak{X}$ of $\mathcal{X}$ which is not contained in $\mathfrak{G}$. In other words, this property of $\mathfrak{H}$ is characteristic for the possibility of suppressing $\mathfrak{H}$ from the map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$ so as to obtain a map $m(\mathfrak{R}, \mathfrak{T}, \Upsilon)$.*

We consider the correspondence $\mathfrak{X}\rightarrow\mathfrak{Y}=\mathfrak{H}'\mathfrak{X}$. Since there is no $\mathfrak{X}$ for which $\mathfrak{H}'\mathfrak{X}$ is void, this correspondence carries $\mathcal{X}$ into Υ in a univocal manner. In order that the correspondence be biunivocal, it is necessary and sufficient that $\mathfrak{H}'\mathfrak{X}_1=\mathfrak{H}'\mathfrak{X}_2$ imply $\mathfrak{X}_1=\mathfrak{X}_2$ whenever $\mathfrak{X}_1$ and $\mathfrak{X}_2$ are members of $\mathcal{X}$. In order to show that this correspondence is continuous, we consider an arbitrary neighborhood in Υ and its antecedent in $\mathcal{X}$. Every open set in $\mathfrak{T}$ is obtained as a set $\mathfrak{G}\mathfrak{H}'$ where $\mathfrak{G}$ is an open set in $\mathfrak{S}$. Hence the neighborhoods in Υ are

specified by the relations $\mathfrak{H}'\mathfrak{X}=\mathfrak{Y}\subset\mathfrak{H}'\mathfrak{G}$, where $\mathfrak{G}$ runs over the class of all open sets in $\mathfrak{S}$. Since each such relation is equivalent to the corresponding relation $\mathfrak{X}\subset\mathfrak{G}\cup\mathfrak{H}$, where $\mathfrak{G}\cup\mathfrak{H}$ is open in $\mathfrak{S}$, we see that each neighborhood in Υ has as its antecedent a neighborhood in $\mathcal{X}$. This property of the correspondence identifies it as a continuous correspondence. In order that the correspondence be a topological equivalence it must be biunivocal and bicontinuous. When the condition of being biunivocal is met in accordance with the criterion noted above, we see therefore that the correspondence is an equivalence if and only if the inverse correspondence $\mathfrak{H}'\mathfrak{X}=\mathfrak{Y}\rightarrow\mathfrak{X}$ is continuous. By the Cauchy criterion for continuity,* the latter condition is satisfied if and only if $\mathfrak{G}\supset\mathfrak{X}_0$, where $\mathfrak{G}$ is open in $\mathfrak{S}$, implies the existence of an open set $\mathfrak{G}_0\mathfrak{H}'$ in $\mathfrak{T}$, $\mathfrak{G}_0$ being open in $\mathfrak{S}$, such that $\mathfrak{Y}_0=\mathfrak{H}'\mathfrak{X}_0$ is contained in $\mathfrak{G}_0\mathfrak{H}'$ while $\mathfrak{Y}=\mathfrak{H}'\mathfrak{X}\subset\mathfrak{G}_0\mathfrak{H}'$ implies $\mathfrak{X}\subset\mathfrak{G}$. This characteristic property is obviously equivalent to the one stated in the theorem. We complete the proof of the theorem by showing that the property stated in the theorem implies the property which was recognized above as sufficient to make the correspondence $\mathfrak{X}\rightarrow\mathfrak{Y}$ biunivocal. Suppose that $\mathfrak{X}_1\neq\mathfrak{X}_2$, the notation being chosen so that some point of $\mathfrak{X}_2$ does not belong to $\mathfrak{X}_1$. Then there exists an open set $\mathfrak{G}$ in $\mathfrak{S}$ which contains $\mathfrak{X}_1$ but not $\mathfrak{X}_2$. By virtue of the property which is now being assumed for the set $\mathfrak{H}$, there exists an open set $\mathfrak{G}_1$ such that $\mathfrak{G}_1\cup\mathfrak{H}$ contains $\mathfrak{X}_1$ but contains no set $\mathfrak{X}$ which is not contained in $\mathfrak{G}$. In particular, then, $\mathfrak{G}_1\cup\mathfrak{H}$ does not contain $\mathfrak{X}_2$. It is therefore true that $\mathfrak{H}'\mathfrak{X}_1\neq\mathfrak{H}'\mathfrak{X}_2$: for otherwise we should have $\mathfrak{H}'\mathfrak{X}_2=\mathfrak{H}'\mathfrak{X}_1\subset\mathfrak{G}\mathfrak{H}'$ and $\mathfrak{X}_2\subset\mathfrak{G}_1\cup\mathfrak{H}$, contrary to fact.

Because of the possibility of suppressing from a given map an open set such as that described in Theorem 18, it is convenient to introduce the following definition.

DEFINITION 9. *A map $m(\mathfrak{R}, \mathfrak{S}, \mathcal{X})$ is said to be redundant if there exists a non-void open set $\mathfrak{H}$ which can be suppressed from $\mathfrak{S}$ in accordance with Theorem* 18; *the set $\mathfrak{H}$ is called a set of redundancy. A map in which no such set exists is said to be irredundant.*

In general, a map will be redundant and will not even be convertible into an irredundant map by single or successive removals of sets of redundancy. We may, however, note two simple positive results in this connection. First, we have

THEOREM 19. *If the map $m(\mathfrak{R}, \mathfrak{S}_1, \mathcal{X}_1)$ is converted into a map $m(\mathfrak{R}, \mathfrak{S}_2, \mathcal{X}_2)$ by the suppression of a set of redundancy $\mathfrak{H}_1$, and if the map $m(\mathfrak{R}, \mathfrak{S}_2, \mathcal{X}_2)$ is converted in its turn into a map $m(\mathfrak{R}, \mathfrak{S}_3, \mathcal{X}_3)$ by the suppression of a set of*

* AH, pp. 53–54, Satz IV.

redundancy $\mathfrak{H}_2$, *then the set* $\mathfrak{H}=\mathfrak{H}_1 \cup \mathfrak{H}_2$ *is a set of redundancy in the map* $m(\mathfrak{R}, \mathfrak{S}_1, \mathfrak{X}_1)$ *and its suppression yields the map* $m(\mathfrak{R}, \mathfrak{S}_3, \mathfrak{X}_3)$.

Since $\mathfrak{S}_2$ is a closed proper subset of $\mathfrak{S}_1$ and $\mathfrak{S}_3$ is a closed proper subset of $\mathfrak{S}_2$, the set $\mathfrak{S}_3$ is a closed proper subset of $\mathfrak{S}_1$ and its complement $\mathfrak{S}_3' = \mathfrak{S}_2' \cup \mathfrak{S}_2\mathfrak{S}_3' = \mathfrak{H}_1 \cup \mathfrak{H}_2 = \mathfrak{H}$ is a non-void open subset of $\mathfrak{S}_1$. The suppression of $\mathfrak{H}$ obviously converts $m(\mathfrak{R}, \mathfrak{S}_1, \mathfrak{X}_1)$ into $m(\mathfrak{R}, \mathfrak{S}_3, \mathfrak{X}_3)$. Hence $\mathfrak{H}$ is a set of redundancy in the first map.

THEOREM 20. *If* $\mathfrak{X}$ *is densely distributed in* $\mathfrak{S}$, *the map* $m(\mathfrak{R}, \mathfrak{S}, \mathfrak{X})$ *is irredundant.*

The proof is obvious.

A further possibility of simplification in the theory of maps is that of replacing the neighborhood-system described in Theorem 14 by a suitable equivalent subsystem. In this connection we have the following result.

THEOREM 21. *If the sets* $\mathfrak{x}$ *in* $\mathfrak{X}$ *are bicompact, then the subfamilies of* $\mathfrak{X}$ *specified by the relations* $\mathfrak{x} \subset \mathfrak{G}$ *constitute a basis in the* T_0*-space* $\mathfrak{X}$ *even when the set* $\mathfrak{G}$ *is restricted to be a finite union of open sets belonging to an arbitrary fixed basis in* $\mathfrak{S}$. *The character of the* T_0*-space* $\mathfrak{X}$ *then does not exceed the character of* $\mathfrak{S}$, *if the latter is infinite.*

We have to show that, if $\mathfrak{G}$ is any open set in $\mathfrak{S}$ and $\mathfrak{x}$ any set in $\mathfrak{X}$ which is contained in $\mathfrak{G}$, then there exists a set $\mathfrak{G}_1$ of the indicated type for which $\mathfrak{x} \subset \mathfrak{G}_1 \subset \mathfrak{G}$. If $\mathfrak{s}$ is any point in $\mathfrak{x}$, then there exists an open set $\mathfrak{G}(\mathfrak{s})$ which belongs to the chosen basis in $\mathfrak{S}$ and which has the property $\mathfrak{s} \in \mathfrak{G}(\mathfrak{s}) \subset \mathfrak{G}$. Since the sets $\mathfrak{G}(\mathfrak{s})$, $\mathfrak{s} \in \mathfrak{x}$, cover the bicompact set $\mathfrak{x}$, there exist points $\mathfrak{s}_1, \cdots, \mathfrak{s}_n$ in $\mathfrak{x}$ such that $\mathfrak{x} \subset \mathfrak{G}(\mathfrak{s}_1) \cup \cdots \cup \mathfrak{G}(\mathfrak{s}_n) \subset \mathfrak{G}$. We may therefore put $\mathfrak{G}_1 = \mathfrak{G}(\mathfrak{s}_1) \cup \cdots \cup \mathfrak{G}(\mathfrak{s}_n)$. Since the cardinal number of the class of open sets described in the theorem is equal to that of the chosen basis in $\mathfrak{S}$, when the latter is infinite, we see that the character of $\mathfrak{X}$ has the property asserted above.

We come finally to our third problem. Here we can give a complete solution, due originally to Kolmogoroff.* It is embodied in the following theorem.

THEOREM 22. *In order that the correspondence from* $\mathfrak{S}(\mathfrak{X})$ *to* $\mathfrak{X}$ *given by* $\mathfrak{s} \rightarrow \mathfrak{x}$ *where* $\mathfrak{s} \in \mathfrak{x}$ *be univocal and continuous, it is necessary and sufficient that the family* $\mathfrak{X}$ *be continuous. Hence the map* $m(\mathfrak{R}, \mathfrak{S}, \mathfrak{X})$ *characterizes* $\mathfrak{R}$ *as a continuous image of* $\mathfrak{S}(\mathfrak{X})$ *if and only if the family* $\mathfrak{X}$ *is continuous.*

A necessary and sufficient condition that the given correspondence be univocal is clearly that the sets belonging to $\mathfrak{X}$ be disjoint. When this condition is fulfilled, the continuity of the family $\mathfrak{X}$ is known to be both necessary

* AH, p. 61, footnote.

and sufficient for the continuity of the correspondence from the space $\mathfrak{S}(\mathfrak{X})$ to the T_0-space $\mathfrak{X}$, by virtue of results due to Kolmogoroff.* Hence we complete the proof of the present theorem by showing that in a continuous family the distinct members are disjoint. If $\mathfrak{X}_1 \neq \mathfrak{X}_2$ we suppose that the notation is so chosen that $\mathfrak{X}_2$ contains a point $\mathfrak{s}$ which does not belong to $\mathfrak{X}_1$. Then there exists an open set $\mathfrak{G}$ in $\mathfrak{S}$ which contains $\mathfrak{X}_1$ but neither $\mathfrak{s}$ nor $\mathfrak{X}_2$. By the continuity of the family $\mathfrak{X}$, there exists also an open set $\mathfrak{G}_1$ which is contained in $\mathfrak{G}$ and contains $\mathfrak{X}_1$ while every set $\mathfrak{X}$ such that $\mathfrak{X}\mathfrak{G}_1 \neq 0$ is contained in $\mathfrak{G}$. Hence in particular we have $\mathfrak{X}_2\mathfrak{G}_1 = 0$, $\mathfrak{X}_2 \subset \mathfrak{G}_1'$. Since $\mathfrak{X}_1 \subset \mathfrak{G}_1$, the sets $\mathfrak{X}_1$ and $\mathfrak{X}_2$ are disjoint.

With this, we have completed a general survey of the theory of maps. In concluding the present section, we shall summarize the preceding results as they apply to Boolean maps. We obtain the following theorem.

THEOREM 23. *If $\mathfrak{X}$ is an arbitrary family of distinct closed subsets $\mathfrak{X}$ of a bicompact Boolean space $\mathfrak{B}$, the topology introduced in $\mathfrak{X}$ by assigning each non-void subfamily of $\mathfrak{X}$ specified by a relation $\mathfrak{X} \subset \mathfrak{G}$, where $\mathfrak{G}$ is open in $\mathfrak{B}$, as a neighborhood of every one of its elements is equivalent to that obtained by restricting the sets $\mathfrak{G}$ to be bicompact as well as open. Since the closed set $\mathfrak{S}^{-}(\mathfrak{X})$ is a bicompact Boolean space, it may be assumed without loss of generality that $\mathfrak{S}^{-}(\mathfrak{X}) = \mathfrak{B}$. Then the topological space $\mathfrak{X}$ has the properties:*

(1) *it is a T_0-space, of character not exceeding that of $\mathfrak{B}$ if $\mathfrak{B}$ is infinite;*

(2) *it is a T_1-space if and only if no member of the family $\mathfrak{X}$ contains another as a proper subset;*

(3) *it is an H-space if the distinct members of $\mathfrak{X}$ are disjoint.*

A Boolean map $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ exists only if $\mathfrak{R}$ has the properties necessitated by (1), (2), *and* (3). *The suppression of a set of redundancy from a Boolean map yields a Boolean map.*

Since the bicompact open subsets constitute a basis in $\mathfrak{B}$ in accordance with Theorems 1 and 2, the first statement of the present theorem is justified by reference to Theorem 21. The second statement follows from Theorems 3 and 16. The properties (1), (2), and (3) are established by reference to Theorems 14, 17, and 21. If $\mathfrak{H}$ is a set of redundancy in the map $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$, then $\mathfrak{H}'$ is a bicompact Boolean subspace of $\mathfrak{B}$ in accordance with Theorem 3; it follows that the suppression of $\mathfrak{H}$ yields a Boolean map.

2. **Construction of Boolean maps.** In this section, we propose to show that the bicompact Boolean spaces are universal mapping spaces; in other words, that every T_0-space can be mapped in an appropriate bicompact Boolean space. The constructive method which we employ is naturally

* Compare AH, p. 67, Satz VI.

algebraic in character. Once the problem of finding some special type of universal mapping space has been proposed, the bicompact Boolean spaces appear to be peculiarly fitted to provide a solution. Thus the property of total-disconnectedness seems to be desirable if we are to avoid the consequences of the theorem which states that every continuous image of a connected space is connected;* and the property of bicompactness seems equally desirable if we are to recover from the theory of maps some of the known theorems concerning the representation of bicompact H-spaces.† Moreover, if we restrict attention for the moment to separable spaces, that is, to spaces of character not exceeding $\aleph_0$, we have positive evidence in favor of our selection of Boolean spaces as universal mapping spaces. For it is known that every compact metric space or, equivalently, every bicompact separable H-space is a continuous image of the Cantor discontinuum $\mathfrak{D}$ and hence of the equivalent Boolean space $\mathfrak{B}_{\mathfrak{c}}$, $\mathfrak{c}=\aleph_0$, of Theorems 9 and 13.‡

In order to give a reasonably complete analysis of the theory of Boolean maps, we shall have to proceed in a somewhat more complicated way than would be necessary if we were concerned merely with solving the problem of universality raised in the preceding paragraph. We begin therefore with some propositions about topological algebra.

THEOREM 24. *Let* $\mathfrak{R}$ *be an arbitrary* T_0*-space,* $B_{\mathfrak{R}}$ *the Boolean ring of all subsets of* $\mathfrak{R}$*,* $\mathfrak{a}_{\mathfrak{R}}$ *the class of all nowhere dense subsets of* $\mathfrak{R}$*. Then* $\mathfrak{a}_{\mathfrak{R}}$ *is an ideal in* $B_{\mathfrak{R}}$*; and, if* A *is any subring of* $B_{\mathfrak{R}}$*, the class* $\mathfrak{a}$ *of all elements common to* $\mathfrak{a}_{\mathfrak{R}}$ *and* A *is an ideal in* A*. The subclass* $A_{\mathfrak{R}}$ *of* $B_{\mathfrak{R}}$ *specified by any of the following equivalent conditions:*

(1) $a^-a'\epsilon\mathfrak{a}_{\mathfrak{R}}$; (2) $aa'^-\epsilon\mathfrak{a}_{\mathfrak{R}}$; (3) $a^-a'^-\epsilon\mathfrak{a}_{\mathfrak{R}}$;

(4) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to an open subset of* $\mathfrak{R}$;

(5) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to a regular open subset of* $\mathfrak{R}$;

(6) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to a closed subset of* $\mathfrak{R}$;

(7) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to its interior* $a'^{-'}$;

(8) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to its closure* a^-;

(9) *a is congruent* (mod $\mathfrak{a}_{\mathfrak{R}}$) *to the regular open set* $a^{-'-'}$;

is a subring of $B_{\mathfrak{R}}$ *containing the set* $\mathfrak{R}$ *as its unit and the class* $\mathfrak{a}_{\mathfrak{R}}$ *as an ideal. It is the subring of* $B_{\mathfrak{R}}$ *generated by* $\mathfrak{a}_{\mathfrak{R}}$ *and the class of all open* (*regular open, closed*) *subsets of* $\mathfrak{R}$.

By definition the members of $\mathfrak{a}_{\mathfrak{R}}$ are characterized by the identity $a^{-'-}=\mathfrak{R}$. To prove that $\mathfrak{a}_{\mathfrak{R}}$ is an ideal, we combine this characteristic property

* AH, p. 53, Corollary to Satz I.
† AH, pp. 96–98.
‡ AH, pp. 85–88, 119–122.

with R Theorem 16. Thus if a and b are nowhere dense, we have $\mathfrak{R}=a^{-\prime-}$ $=(a^{-\prime}\mathfrak{R})^{-}=(a^{-\prime}(b^{-}\vee b^{-\prime}))^{-}=(a^{-\prime}b^{-}\vee a^{-\prime}b^{-\prime})^{-}=(a^{-\prime}b^{-})^{-}\vee(a^{-\prime}b^{-\prime})^{-}<a^{-\prime-}b^{-}$ $\vee(a^{-\prime}b^{-\prime})^{-}=b^{-}\vee(a^{-\prime}b^{-\prime})^{-}$, $(a^{-\prime}b^{-\prime})^{-}>b^{-\prime}$, $(a^{-\prime}b^{-\prime})^{-}>b^{-\prime-}=\mathfrak{R}$, and $(a\vee b)^{-\prime-}$ $=(a^{-}\vee b^{-})^{\prime-}=(a^{-\prime}b^{-\prime})^{-}=\mathfrak{R}$, so that $a\vee b$ is also nowhere dense. Similarly, if a is nowhere dense and $c<a$, then $a^{-}>c^{-}$, $c^{-\prime}>a^{-\prime}$, and $c^{-\prime-}>a^{-\prime-}=\mathfrak{R}$, so that c is nowhere dense. The void set obviously belongs to $\mathfrak{a}_{\mathfrak{R}}$. It is thus clear that $\mathfrak{a}_{\mathfrak{R}}$ is an ideal in $B_{\mathfrak{R}}$; and also that in any subring A of $B_{\mathfrak{R}}$ the class $\mathfrak{a}$ of elements common to $\mathfrak{a}_{\mathfrak{R}}$ and A is an ideal.

We shall next consider the class $A_{\mathfrak{R}}$ specified by the condition (3) above. First we observe that $A_{\mathfrak{R}}$ contains $\mathfrak{a}_{\mathfrak{R}}$: for, if a is nowhere dense, then a^{-} and $a^{-}a^{\prime-}<a^{-}$ are nowhere dense. Then it is obvious from the symmetry of the condition (3) in a and a' that $A_{\mathfrak{R}}$ contains a if and only if it contains a'. Furthermore, $A_{\mathfrak{R}}$ contains $a\vee b$ together with a and b: for the relations $(a\vee b)^{-}=a^{-}\vee b^{-}$, $(a\vee b)^{\prime-}=(a'b')^{-}<a^{\prime-}b^{\prime-}$ imply $(a\vee b)^{-}(a\vee b)^{\prime-}<a^{-}a^{\prime-}$ $\vee\, b^{-}b^{\prime-}\epsilon\,\mathfrak{a}_{\mathfrak{R}}$ when a and b are in $A_{\mathfrak{R}}$. Hence we see that $A_{\mathfrak{R}}$ is a subring of $B_{\mathfrak{R}}$, since it contains $ab=(a'\vee b')'$ and $a+b=ab'\vee a'b=(a'\vee b)'\vee(a\vee b')'$ whenever it contains a and b. The previous results show that $\mathfrak{a}_{\mathfrak{R}}$ is an ideal in $A_{\mathfrak{R}}$.

Now if a is any open set in $\mathfrak{R}$, we have $(a^{-}a^{\prime-})^{-\prime-}=(a^{-}a')^{-\prime-}=(a^{-}a')^{\prime-}$ $=(a^{-\prime}\vee a)^{-}=a^{-\prime-}\vee a^{-}>a^{-\prime}\vee a^{-}=\mathfrak{R}$, so that $a^{-}a^{\prime-}\epsilon\,\mathfrak{a}_{\mathfrak{R}}$, $a\,\epsilon A_{\mathfrak{R}}$; and, if $a\equiv b \pmod{\mathfrak{a}_{\mathfrak{R}}}$ where b is an open set in $\mathfrak{R}$, we have $a=(a+b)+b$ where $a+b\,\epsilon\,\mathfrak{a}_{\mathfrak{R}}\subset A_{\mathfrak{R}}$, $b\,\epsilon A_{\mathfrak{R}}$ and hence $a\,\epsilon A_{\mathfrak{R}}$. Thus (4) implies (3). Since the relation $a\equiv b \pmod{\mathfrak{a}_{\mathfrak{R}}}$ is equivalent to $a'\equiv b' \pmod{\mathfrak{a}_{\mathfrak{R}}}$, we see immediately that (6) also implies (3). It is trivial, from the preceding results, that (5) implies (3). Next we see by virtue of the relations $a^{-}a'<a^{-}a^{\prime-}$, $aa^{\prime-}<a^{-}a^{\prime-}$ that (3) implies both (1) and (2). Now (1) obviously implies $a\equiv a^{-} \pmod{\mathfrak{a}_{\mathfrak{R}}}$ since $a^{-}+a=a^{-}a'\vee a^{-\prime}a=a^{-}a'\epsilon\,\mathfrak{a}_{\mathfrak{R}}$; and (2) similarly implies $a\equiv a^{\prime-\prime} \pmod{\mathfrak{a}_{\mathfrak{R}}}$. Thus (1) implies (8), (2) implies (7). It is evident that (8) implies (6) and that (7) implies (4). From the scheme of implications

$$(3)\begin{array}{l}\nearrow (1)\rightarrow(8)\rightarrow(6)\searrow\\ \searrow (2)\rightarrow(7)\rightarrow(4)\nearrow\end{array}(3)$$

we see that the conditions (1)–(4), (6)–(8) are equivalent. Since (3) implies $a\equiv a^{-} \pmod{\mathfrak{a}_{\mathfrak{R}}}$, since $a\equiv a^{-} \pmod{\mathfrak{a}_{\mathfrak{R}}}$ implies $a^{-}\epsilon A_{\mathfrak{R}}$, and since (3) then implies $a^{-}\equiv a^{-\prime-\prime} \pmod{\mathfrak{a}_{\mathfrak{R}}}$, we conclude that (3) implies (9) $a\equiv a^{-\prime-\prime} \pmod{\mathfrak{a}_{\mathfrak{R}}}$. It is obvious that (9) implies (5). Thus by fitting the scheme of implications

$$(3)\rightarrow(9)\rightarrow(5)\rightarrow(4)$$

into the scheme given above, we see that the conditions (1)–(9) are equivalent. It is then obvious that $A_{\mathfrak{R}}$ is the subring of $B_{\mathfrak{R}}$ generated in the manner described above.

We may remark that some of the conditions (1)–(9) may be cast into more geometrical form. Thus (1), (2), (3) assert respectively that the border, the frontier, or the boundary of a is nowhere dense; (7) asserts that a differs from its interior by a nowhere dense set; and (8) asserts that a differs from its closure by a nowhere dense set.

While the regular open sets have little interest for us in the present chapter, they will later play an important rôle. Hence we shall enumerate some of their useful properties at this point.

THEOREM 25. *Between the residual classes* (mod $\mathfrak{a}_{\mathfrak{R}}$) *in the Boolean ring* $A_{\mathfrak{R}}$ *and the regular open sets in* $\mathfrak{R}$, *there exists a biunivocal correspondence such that each residual class contains the corresponding regular open set. The intersection of a finite number of regular open sets is a regular open set; but the union, in general, is not. If* $a \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$) *where* b *is a regular open set, then the interior of* a *is contained in* b; *and, if* a *is an arbitrary subset of a regular open set* b, *then* $a^{-\prime-\prime}$ *is a regular open set contained in* b. *The regular open sets in* $\mathfrak{R}$ *are characterized by the property that their borders coincide with the borders of their exteriors.*

From the preceding theorem, we know that each residual class (mod $\mathfrak{a}_{\mathfrak{R}}$) contains at least one regular open set. Thus we have to prove that, if a and b are both regular open sets, then $a \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$) implies $a = b$. Instead of establishing this result directly, we first show that, if a and b are both regular open sets, then their intersection is also such a set. We have $(ab)^- < a^- b^-$, $(ab)^{-\prime} > a^{-\prime} \vee b^{-\prime}$, $(ab)^{-\prime-} > a^{-\prime-} \vee b^{-\prime-}$, $(ab)^{-\prime-\prime} < a^{-\prime-\prime} b^{-\prime-\prime} = ab$. On the other hand, if c is an arbitrary set we have $c < c^-$, $c' > c^{-\prime}$, $c'^- > c^{-\prime-}$, and $c'^{-\prime} < c^{-\prime-\prime}$. In particular, since ab is open, we have $ab = (ab)'^{-\prime} < (ab)^{-\prime-\prime}$. We therefore conclude that $ab = (ab)^{-\prime-\prime}$, as we wished to do. It is easy to see by examples that the union of regular open sets need not be a regular open set. We can now return to the previous question. To begin with, let us assume that the regular open sets a and b satisfy the relation $a > b$ in addition to the relation $a \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$). Then we see that $ab' = a + b \epsilon \mathfrak{a}_{\mathfrak{R}}$, $(ab')^{-\prime-} = \mathfrak{R}$. Hence we have $(ab')^- > ab'$, $(ab')^{-\prime} < a' \vee b$, $\mathfrak{R} = (ab')^{-\prime-} < a'^- \vee b^- = a' \vee b^-$, $b^- > a$, $b^{-\prime} < a'$, $b^{-\prime-} < a'^- = a'$, $b = b^{-\prime-\prime} > a$, and $a = b$. Passing now to the general case, we see that ab is a regular open set satisfying the relations $a \equiv aa \equiv ab \equiv bb \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$) and the relations $a > ab$, $b > ab$. Thus the result just obtained implies $a = ab = b$, as we wished to show. Now if $a \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$) where b is a regular open set, we know from Theorem 24 that $a \equiv a'^{-\prime} \equiv a^{-\prime-\prime} \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$); and from the results proved above we know that $a'^{-\prime} < a^{-\prime-\prime}$ and that the regular open set $a^{-\prime-\prime}$ coincides with b. Hence we see that the interior of a is contained in b. We may state this result in the form: the modification of a regular

open set by the addition (mod 2) of a nowhere dense set may suppress interior points but cannot adjoin interior points. Next we consider the relations $a<b$, $b=b^{-\prime-\prime}$: we then have $a^-<b^-$, $a^{-\prime}>b^{-\prime}$, $a^{-\prime-}>b^{-\prime-}$, $a^{-\prime-\prime}<b^{-\prime-\prime}=b$, as stated above. Finally, we establish the characterization of the regular open sets given in the theorem. We have to prove that the relations $a=a^{-\prime-\prime}$ and $a^-a'=a^{-\prime-}a^-$ are equivalent. If $a=a^{-\prime-\prime}$, we have $a'=a^{-\prime-}$ and hence $a^-a'=a^{-\prime-}a^-$ as asserted. If $a^-a'=a^{-\prime-}a^-$, we have $a^{-\prime}\vee a=a^{-\prime-\prime}\vee a^{-\prime}$, $a=a^-(a^{-\prime}\vee a)=a^-(a^{-\prime-\prime}\vee a^{-\prime})=a^-a^{-\prime-\prime}$, where $a^->a^{-\prime-\prime}$ by virtue of the relation $a^{-\prime}<a^{-\prime-}$; and we therefore conclude that $a=a^{-\prime-\prime}$.

We now proceed to the algebraic construction of Boolean maps.

THEOREM 26. *Let $\mathfrak{R}$ be any non-void T_0-space, $A_{\mathfrak{R}}$ the Boolean ring of Theorem* 24, *and A any subring of $A_{\mathfrak{R}}$ with the properties*

(1) *A contains the set $\mathfrak{R}$ as its unit;*

(2) *the interiors of sets in A constitute a basis for $\mathfrak{R}$.*

Then let $\mathfrak{B}=\mathfrak{E}(A)$ be the representative bicompact Boolean space for the ring A; let $\mathfrak{a}(\mathfrak{r})$ be the class of all those sets a in A such that a^- does not contain the point $\mathfrak{r}$ in $\mathfrak{R}$; and let $\mathfrak{b}(\mathfrak{r})$ be the class of all those sets b in A such that $\mathfrak{r}$ is interior to b. The class $\mathfrak{a}(\mathfrak{r})$ is an ideal in A; and the closed set $\mathfrak{X}(\mathfrak{r})=\mathfrak{E}'(\mathfrak{a}(\mathfrak{r}))$ in $\mathfrak{E}(A)$ has the property

$$\mathfrak{X}(\mathfrak{r}) = \Big(\sum_{a\epsilon\mathfrak{a}(\mathfrak{r})}\mathfrak{E}(a)\Big)' = \prod_{b\epsilon\mathfrak{b}(\mathfrak{r})}\mathfrak{E}(b).$$

If X is the family of all sets $\mathfrak{X}(\mathfrak{r})$ corresponding to points $\mathfrak{r}$ in $\mathfrak{R}$, the correspondence $\mathfrak{r}\to\mathfrak{X}(\mathfrak{r})$ defines a Boolean map $m(\mathfrak{R},\mathfrak{B},\mathsf{X})$. In order that a Boolean ring A of subsets of $\mathfrak{R}$ have the properties demanded here, it is sufficient that it be the ring with unit $\mathfrak{R}$ generated by an arbitrary basis in $\mathfrak{R}$.

We first prove that $\mathfrak{a}(\mathfrak{r})$ is an ideal. Since $0^-=0$, the void set 0 is in $\mathfrak{a}(\mathfrak{r})$. If a and b are in $\mathfrak{a}(\mathfrak{r})$, the relation $(a\vee b)^-=a^-\vee b^-$ shows that $a\vee b$ is also in $\mathfrak{a}(\mathfrak{r})$; and if $c<a$, where a is in $\mathfrak{a}(\mathfrak{r})$, the relation $c^-<a^-$ shows that c is also in $\mathfrak{a}(\mathfrak{r})$. Thus $\mathfrak{a}(\mathfrak{r})$ is an ideal in accordance with R Theorem 16. Next we consider the relation between the ideal $\mathfrak{a}(\mathfrak{r})$ and the class $\mathfrak{b}(\mathfrak{r})$. We see at once that a is in $\mathfrak{a}(\mathfrak{r})$ if and only if $b=a'$ is in $\mathfrak{b}(\mathfrak{r})$: for $a\epsilon\mathfrak{a}(\mathfrak{r})$ implies $\mathfrak{r}\epsilon a^{-\prime}=b^{\prime-\prime}$ and $b\epsilon\mathfrak{b}(\mathfrak{r})$ implies $\mathfrak{r}\epsilon b^{\prime-\prime}=a^{-\prime}$. It is therefore evident that

$$\mathfrak{X}(\mathfrak{r}) = \mathfrak{E}'(\mathfrak{a}(\mathfrak{r})) = \Big(\sum_{a\epsilon\mathfrak{a}(\mathfrak{r})}\mathfrak{E}(a)\Big)' = \prod_{a\epsilon\mathfrak{a}(\mathfrak{r})}\mathfrak{E}'(a) = \prod_{a\epsilon\mathfrak{a}(\mathfrak{r})}\mathfrak{E}(a') = \prod_{b\epsilon\mathfrak{b}(\mathfrak{r})}\mathfrak{E}(b)$$

by virtue of R Theorem 67. It is therefore easy to determine the significance of the relation $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{E}(b)$ when b is an element of A. It is equivalent successively to the relations $\mathfrak{E}'(\mathfrak{a}(\mathfrak{r}))\subset\mathfrak{E}(b)$, $\mathfrak{E}(\mathfrak{a}(\mathfrak{r}))\supset\mathfrak{E}(b')=\mathfrak{E}'(b)$, $b'\epsilon\mathfrak{a}(\mathfrak{r})$, and

$b\epsilon\mathfrak{b}(\mathfrak{r})$. Thus the relation $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(b)$ holds if and only if $\mathfrak{r}$ is interior to the set b. From this fact we can deduce that the sets $\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{X}(\mathfrak{s})$ corresponding to distinct points $\mathfrak{r}$ and $\mathfrak{s}$ in $\mathfrak{R}$ are necessarily distinct sets. Since the interiors of sets in A constitute a basis in $\mathfrak{R}$, there exists a set b in A to which just one of the two points $\mathfrak{r}$ and $\mathfrak{s}$ is interior. If we suppose that the notation is chosen so that $\mathfrak{r}$ is interior to b while $\mathfrak{s}$ is not, we conclude that $\mathfrak{C}(b)$ contains $\mathfrak{X}(\mathfrak{r})$ but not $\mathfrak{X}(\mathfrak{s})$. Hence we have $\mathfrak{X}(\mathfrak{r}) \neq \mathfrak{X}(\mathfrak{s})$ whenever $\mathfrak{r} \neq \mathfrak{s}$.

We shall now interpret the foregoing algebraic facts topologically. In the bicompact Boolean space $\mathfrak{B} = \mathfrak{C}(A)$, defined as in Theorem 1, the sets $\mathfrak{C}(a)$, $a\epsilon A$, are bicompact open sets constituting a basis; and the sets $\mathfrak{X}(\mathfrak{r})$ are closed subsets of $\mathfrak{B}$ in accordance with Theorems 1 and 4. Since the closed sets $\mathfrak{X}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$ are distinct and obviously non-void, they constitute a family $\mathcal{X}$ to which Theorems 14 and 23 can be applied. The subfamilies of $\mathcal{X}$ specified by the relations $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(b)$, $b\epsilon A$, then constitute a basis in the T_0-space $\mathcal{X}$, as asserted in Theorem 23. The correspondence $\mathfrak{r} \rightarrow \mathfrak{X}(\mathfrak{r})$ carries $\mathfrak{R}$ into $\mathcal{X}$ in a biunivocal way. Since the sets $b'^{-\prime}$, where b is in A, constitute a basis in $\mathfrak{R}$ and since the relations $\mathfrak{r}\epsilon b'^{-\prime}$ and $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(b)$ are equivalent, we now see that the correspondence sets up a topological equivalence between the spaces $\mathcal{X}$ and $\mathfrak{R}$. By Definition 7, this relation between $\mathfrak{R}$, $\mathfrak{B}$, and $\mathcal{X}$ is a Boolean map $m(\mathfrak{R}, \mathfrak{B}, \mathcal{X})$. The closing remark of the theorem is an obvious consequence of Theorem 24; it shows that the existence of a ring A of the required type offers no difficulty.

For convenience of reference to this fundamental result we introduce the following terminology.

Definition 10. *If $\mathfrak{R}$ is a T_0-space and $A_{\mathfrak{R}}$ the Boolean ring described in Theorem* 24, *then any subring A of $A_{\mathfrak{R}}$ with the properties* (1) *and* (2) *of Theorem* 26 *is called a basic ring of $\mathfrak{R}$; the ring $A_{\mathfrak{R}}$ is called the complete basic ring of $\mathfrak{R}$.*

Definition 11. *If $\mathfrak{R}$ is a T_0-space and A is a basic ring of $\mathfrak{R}$, then the Boolean map $m(\mathfrak{R}, \mathfrak{B}, \mathcal{X})$ defined by A in the manner described in Theorem* 26 *is called an algebraic* (*Boolean*) *map of $\mathfrak{R}$; and the map defined by the ring $A_{\mathfrak{R}}$ is called the complete algebraic* (*Boolean*) *map of $\mathfrak{R}$.*

Before proceeding to a more detailed analysis of the algebraic maps, we pause to consider the general significance of Theorems 23 and 26. It is clear that we can summarize the situation in the following terms.

Theorem 27. *The algebraic theory of Boolean rings* (*with unit*) *is mathematically equivalent to the topological theory of T_0-spaces.*

In the first place, Theorem 23 shows that any family $\mathcal{X}$ of distinct non-

void closed sets $\mathfrak{X}$ in a bicompact Boolean space $\mathfrak{B}$ may be regarded, under a suitable topology, as a T_0-space; but, since the open sets $\mathfrak{X}'$ in $\mathfrak{B}$ correspond by Theorems 1 and 4 to ideals in a Boolean ring with unit which has $\mathfrak{B}$ as its representative, we actually have a representation of certain ideals and their algebraic relations by means of the indicated T_0-space and its topological properties. On the other hand Theorem 26 shows that such representative T_0-spaces are entirely arbitrary and unrestricted. Thus the algebraic structure of families of ideals in a Boolean ring with unit is exactly reflected in the topological structure of T_0-spaces. It follows that the complete analysis of the ideal structure of Boolean rings is as complicated as the analysis of the structure of all T_0-spaces. In the second place, Theorems 23 and 26 may be viewed, from another angle, as placing the study of all T_0-spaces on a purely combinatorial basis: for any such space is completely described as a configuration of ideals in a Boolean ring; and, as we have remarked elsewhere, the postulates for Boolean rings are postulates for an abstract algebra of combinations. These theorems, furthermore, reduce the construction of T_0-spaces to a kind of tactical game with the closed subsets of bicompact Boolean spaces.* We may remark that the requirement that the Boolean rings in the foregoing discussion should have units, does not affect the range of our comments in a serious way: for the adjunction of a unit to a Boolean ring is an essentially trivial operation, as we have already seen in R Theorem 1 and Theorem 8.

We now return to the further discussion of algebraic Boolean maps.

THEOREM 28. *If $\mathfrak{a}$ is the ideal of all nowhere dense sets in a basic ring A of a T_0-space $\mathfrak{R}$, then the algebraic map $m(\mathfrak{R}, \mathfrak{B}, \mathcal{X})$ defined by A has the following properties*:

(1) *in order that $\mathfrak{E}(a)$, $a\epsilon A$, contain no set $\mathfrak{X}(\mathfrak{r})$ in $\mathcal{X}$, it is necessary and sufficient that $a\epsilon\mathfrak{a}$;*

(2) *the set $\mathfrak{H}=\mathfrak{E}(\mathfrak{a})$ is the maximal open subset of $\mathfrak{B}$ which contains no set $\mathfrak{X}(\mathfrak{r})$ in $\mathcal{X}$;*

(3) *the points of $\mathfrak{R}$ specified by the relations $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(a)$, $\mathfrak{X}(\mathfrak{r})\mathfrak{E}(a)\neq 0$, where a is in A, constitute the open set $a'^{-\prime}$ and the closed set a^- respectively.*

If $\mathfrak{R}$ is an H-space, then $\mathfrak{r}\neq\mathfrak{s}$ implies $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s})\mathfrak{E}'(\mathfrak{a})=0$; and, if in addition the basic ring A actually contains a basis for $\mathfrak{R}$, then $\mathfrak{r}\neq\mathfrak{s}$ implies $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s})=0$. The sufficient condition of Theorem 23 thus becomes necessary in the latter case.

---

* The tactical method has been extensively used, for example, by Alexandroff and Urysohn, *Mémoire sur les espaces topologiques compacts*, Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Eerste Sectie, Deel XIV, No. 1 (1929). Their constructions are usually carried out in the Euclidean plane, however.

In particular, in order that $\mathfrak{R}$ be an H-space, it is necessary and sufficient that in its complete map $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ the sets in $\mathfrak{X}$ be disjoint.

In the proof of Theorem 26 we have already established the first part of (3): the relations $\mathfrak{r}\epsilon a'^{-\prime}$ and $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a)$ are equivalent. The second part can be proved from the first as follows: the relation $\mathfrak{X}(\mathfrak{r})\mathfrak{C}(a) \neq 0$ holds if and only if $\mathfrak{X}(\mathfrak{r})$ is not contained in $\mathfrak{C}(a') = \mathfrak{C}'(a)$ and is therefore equivalent to the relation $\mathfrak{r}\epsilon((a')'^{-\prime})' = a^{-}$. It is now clear that (3) implies (1): for, in order that the set $\mathfrak{C}(\mathfrak{a})$ contain no set $\mathfrak{X}(\mathfrak{r})$ it is necessary and sufficient that $a'^{-\prime} = 0$; but Theorem 24 permits us to use the relations $a \equiv a'^{-\prime}$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $a'^{-\prime} < a^{-\prime-\prime}$ to show that $a'^{-\prime} = 0$ implies $a \equiv 0$ (mod $\mathfrak{a}_{\mathfrak{R}}$) or $a\epsilon\mathfrak{a}$ and that $a\epsilon\mathfrak{a}$ implies $a'^{-\prime} < (a^{-\prime -})' = \mathfrak{R}' = 0$. We now show that (1) implies (2). If $\mathfrak{X}(\mathfrak{r})$ is contained in $\mathfrak{C}(\mathfrak{a})$, the open sets $\mathfrak{C}(a)$, $a\epsilon\mathfrak{a}$, cover $\mathfrak{X}(\mathfrak{r})$; since $\mathfrak{X}(\mathfrak{r})$ is bicompact, there exist elements $a_1, \cdots, a_n$ such that $\mathfrak{C}(a) = \mathfrak{C}(a_1) \cup \cdots \cup \mathfrak{C}(a_n) \supset \mathfrak{X}(\mathfrak{r})$, where $a_1, \cdots, a_n$ and $a = a_1 \vee \cdots \vee a_n$ belong to the ideal $\mathfrak{a}$; and consequently we have a contradiction. Thus $\mathfrak{C}(\mathfrak{a})$ contains no set $\mathfrak{X}(\mathfrak{r})$. Since every open set in $\mathfrak{B} = \mathfrak{C}(A)$ is a union of sets $\mathfrak{C}(a)$, $a\epsilon A$, an open set which contains no set $\mathfrak{X}(\mathfrak{r})$ must be a union of sets $\mathfrak{C}(a)$, $a\epsilon\mathfrak{a}$, and must therefore be contained in $\mathfrak{C}(\mathfrak{a})$. Hence $\mathfrak{C}(\mathfrak{a})$ is the maximal open set which contains no set $\mathfrak{X}(\mathfrak{r})$ in $\mathfrak{X}$.

We shall now prove that, if $\mathfrak{R}$ is an H-space, the relation $\mathfrak{r} \neq \mathfrak{s}$ implies $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s})\mathfrak{C}'(\mathfrak{a}) = 0$. By hypothesis, the basic ring A contains elements a and b such that $\mathfrak{r}\epsilon a'^{-\prime}$, $\mathfrak{s}\epsilon b'^{-\prime}$, and $a'^{-\prime}b'^{-\prime} = 0$. By Theorem 24, $ab \equiv a'^{-\prime}b'^{-\prime} \equiv 0$ (mod $\mathfrak{a}_{\mathfrak{R}}$) so that $ab\epsilon\mathfrak{a}$. Hence we have $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s}) \subset \mathfrak{C}(a)\mathfrak{C}(b) = \mathfrak{C}(ab) \subset \mathfrak{C}(\mathfrak{a})$ by (3); and we conclude that $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s})\mathfrak{C}'(\mathfrak{a}) = 0$. In case A contains a basis for $\mathfrak{R}$, we can select elements a and b which are open sets in $\mathfrak{R}$ and which have the properties $\mathfrak{r}\epsilon a = a'^{-\prime}$, $\mathfrak{s}\epsilon b = b'^{-\prime}$, $ab = 0$. We find that $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s}) = 0$. The proof of the theorem is thereby completed.

It is of considerable interest to compare the different algebraic maps of a given T_0-space $\mathfrak{R}$ and to examine the effect of suppressing the set $\mathfrak{H} = \mathfrak{C}(\mathfrak{a})$ from a given algebraic map of $\mathfrak{R}$. We obtain the following result.

THEOREM 29. *Let $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{R}}), \mathfrak{X})$ be the complete algebraic map of a T_0-space $\mathfrak{R}$; let $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X}^A)$ be an arbitrary algebraic map of $\mathfrak{R}$ defined by a basic ring A; let $\mathfrak{a}_{\mathfrak{R}}$ and $\mathfrak{a}$ be the ideals of nowhere dense sets in $A_{\mathfrak{R}}$ and in A respectively; and let $\mathfrak{p} = f(\mathfrak{q})$ be the function defining the bicontinuous univocal correspondence of $\mathfrak{C}(A)$ with $\mathfrak{C}(A_{\mathfrak{R}})$ in accordance with the fact that A is a subring of $A_{\mathfrak{R}}$ with the same unit $\mathfrak{R}$. Then the indicated correspondence has the properties:*

(1) $f^{-1}(\mathfrak{C}(\mathfrak{a})) \subset \mathfrak{C}(\mathfrak{a}_{\mathfrak{R}})$, $f^{-1}(\mathfrak{C}'(\mathfrak{a})) \supset \mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$, $f(\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})) = \mathfrak{C}'(\mathfrak{a})$;

(2) *if* $\mathfrak{Z}^A(\mathfrak{r}) = f^{-1}(\mathfrak{X}^A(\mathfrak{r}))$, *then* $\mathfrak{Z}^A(\mathfrak{r}) \supset \mathfrak{X}(\mathfrak{r})$, $f(\mathfrak{X}(\mathfrak{r})) = f(\mathfrak{Z}^A(\mathfrak{r})) = \mathfrak{X}^A(\mathfrak{r})$;

(3) *if* $\mathfrak{Y}(\mathfrak{r}) = \mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ *and* $\mathfrak{Y}^A(\mathfrak{r}) = \mathfrak{X}^A(\mathfrak{r})\mathfrak{C}'(\mathfrak{a})$, *then* $f^{-1}(\mathfrak{Y}^A(\mathfrak{r}))\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{Z}^A(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{Y}(\mathfrak{r})$, $f(\mathfrak{Y}(\mathfrak{r})) = \mathfrak{Y}^A(\mathfrak{r})$.

The sets $\mathfrak{Y}^A(\mathfrak{r})$ *constitute a family* Υ^A *of non-void closed subsets densely distributed in the bicompact Boolean space* $\mathfrak{E}'(\mathfrak{a})$. *Under the topology of Theorem* 14, *the* T_0*-spaces* Υ^A *corresponding to different basic rings* A *are all topologically equivalent; in particular the space* Υ *corresponding to the ring* $A=A_{\mathfrak{R}}$ *and the space* Υ^A *corresponding to an arbitrary basic ring* A *are topologically equivalent by virtue of the correspondence* $\mathfrak{Y}(\mathfrak{r})\longleftrightarrow\mathfrak{Y}^A(\mathfrak{r})$. *If* $\mathfrak{R}^*$ *is any* T_0*-space topologically equivalent to the spaces* Υ^A, *then* $\mathfrak{R}^*$ *is a continuous image of* $\mathfrak{R}$. *When* $\mathfrak{R}$ *has infinite character, the character of* $\mathfrak{R}^*$ *does not exceed that of* $\mathfrak{R}$. *When* $\mathfrak{R}$ *is an H-space,* $\mathfrak{R}^*$ *is also an H-space; and* $\mathfrak{R}^*$ *is a biunivocal continuous image of* $\mathfrak{R}$. *The suppression of the open set* $\mathfrak{E}(\mathfrak{a})$ *from the algebraic map* $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X}^A)$ *yields an irredundant Boolean map* $m(\mathfrak{R}^*, \mathfrak{E}'(\mathfrak{a}), \Upsilon^A)$.

Since A is a subring of $A_{\mathfrak{R}}$ by hypothesis, Theorem 7 establishes the existence of a continuous univocal correspondence from $\mathfrak{E}(A_{\mathfrak{R}})$ to $\mathfrak{E}(A)$. Since $\mathfrak{E}(A_{\mathfrak{R}})$ is bicompact, the correspondence is necessarily bicontinuous.† The correspondence was so defined that, if the element a in A be considered as an element $a_{\mathfrak{R}}=a$ of $A_{\mathfrak{R}}$, then $f^{-1}(\mathfrak{E}(a))=\mathfrak{E}(a_{\mathfrak{R}})$. Since $a\epsilon\mathfrak{a}$ if and only if $a=a_{\mathfrak{R}}\epsilon\mathfrak{a}_{\mathfrak{R}}$ and $a=a_{\mathfrak{R}}\epsilon A$, we have

$$f^{-1}(\mathfrak{E}(\mathfrak{a})) = f^{-1}\left(\sum_{a\epsilon\mathfrak{a}}\mathfrak{E}(a)\right) = \sum_{a\epsilon\mathfrak{a}} f^{-1}(\mathfrak{E}(a)) = \sum_{a_{\mathfrak{R}}\epsilon\mathfrak{a}_{\mathfrak{R}}}\mathfrak{E}(a_{\mathfrak{R}}) \subset \mathfrak{E}(\mathfrak{a}_{\mathfrak{R}}),$$

and hence also $f^{-1}(\mathfrak{E}'(\mathfrak{a}))\supset\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$. The sets $f^{-1}(\mathfrak{E}(\mathfrak{a}))$, $f^{-1}(\mathfrak{E}'(\mathfrak{a}))$ are respectively open and closed because of the continuity of f. If we now make use of the bicontinuity of f, we see that $f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))$ is a closed set contained in $\mathfrak{E}'(\mathfrak{a})$. If it does not coincide with $\mathfrak{E}'(\mathfrak{a})$, there exists an element a which belongs to A but not to $\mathfrak{a}$ and which has the property that $\mathfrak{E}(a)f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))=0$. By Theorem 28, the relation $\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})\neq 0$ implies the existence of a point $\mathfrak{r}$ in $\mathfrak{R}$ such that $\mathfrak{X}^A(\mathfrak{r})\subset\mathfrak{E}(a)$ or, equivalently, $\mathfrak{r}\epsilon a'^{-\prime}$. Interpreting the latter relation in terms of the map $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{R}}), \mathfrak{X})$ we see that the element $a=a_{\mathfrak{R}}$ in $A_{\mathfrak{R}}$ has the property $\mathfrak{E}(a_{\mathfrak{R}})\supset\mathfrak{X}(\mathfrak{r})$. Hence the set $f^{-1}(\mathfrak{E}(a))=\mathfrak{E}(a_{\mathfrak{R}})$ has a point in common with $\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ in accordance with Theorem 28. Since $f(\mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))\subset f(\mathfrak{E}(a_{\mathfrak{R}}))f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))=\mathfrak{E}(a)f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))$, we reach the contradiction that $\mathfrak{E}(a)f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))\neq 0$. Hence we must have $f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}))=\mathfrak{E}'(\mathfrak{a})$. This completes the proof of (1) above. Since the relations $\mathfrak{r}\epsilon b'^{-\prime}$, $\mathfrak{X}^A(\mathfrak{r})\subset\mathfrak{E}(b)$, $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{E}(b_{\mathfrak{R}})$ for $b_{\mathfrak{R}}=b\epsilon A$ are equivalent, we have

$$\mathfrak{Z}^A(\mathfrak{r}) = f^{-1}(\mathfrak{X}^A(\mathfrak{r})) = f^{-1}\left(\prod_{b\epsilon\mathfrak{b}(\mathfrak{r})}\mathfrak{E}(b)\right) = \prod_{b\epsilon\mathfrak{b}(\mathfrak{r})} f^{-1}(\mathfrak{E}(b)) = \prod_{b_{\mathfrak{R}}\epsilon\mathfrak{b}(\mathfrak{r})}\mathfrak{E}(b_{\mathfrak{R}})\supset\mathfrak{X}(\mathfrak{r})$$

† AH, p. 95, Satz II.

by virtue of Theorem 26. If we make use of the bicontinuity of the correspondence f, we see that $f(\mathfrak{X}(\mathfrak{r}))$ is a closed subset of $\mathfrak{X}^A(\mathfrak{r})$. If it does not coincide with $\mathfrak{X}^A(\mathfrak{r})$, then there exists a set b in A such that $\mathfrak{E}(b)$ contains $f(\mathfrak{X}(\mathfrak{r}))$ but not $\mathfrak{X}^A(\mathfrak{r})$. If we put $b = b_{\mathfrak{R}} \epsilon A$, we see that $\mathfrak{E}(b_{\mathfrak{R}}) = f^{-1}(\mathfrak{E}(b)) \supset \mathfrak{X}(\mathfrak{r})$ and conclude that $\mathfrak{r} \epsilon b_{\mathfrak{R}}'^{-'} = b'^{-'}$. On the other hand, the fact that $\mathfrak{E}(b)$ does not contain $\mathfrak{X}^A(\mathfrak{r})$ leads to the contradiction that $\mathfrak{r}$ does not belong to $b'^{-'}$. We therefore conclude that $f(\mathfrak{X}(\mathfrak{r})) = \mathfrak{X}^A(\mathfrak{r})$. By definition we have $f(\mathfrak{Z}^A(\mathfrak{r})) = \mathfrak{X}^A(\mathfrak{r})$. This completes the proof of (2) above. From the previous results, we know that

$$\begin{aligned} f^{-1}(\mathfrak{Y}^A(\mathfrak{r}))\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) &= f^{-1}(\mathfrak{X}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}))\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) = f^{-1}(\mathfrak{X}^A(\mathfrak{r}))f^{-1}(\mathfrak{E}'(\mathfrak{a}))\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \\ &= \mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \supset \mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{Y}(\mathfrak{r}), \\ f(\mathfrak{Y}(\mathfrak{r})) &= f(\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})) \subset f(\mathfrak{X}(\mathfrak{r}))f(\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})) = \mathfrak{X}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}) = \mathfrak{Y}^A(\mathfrak{r}), \\ f(f^{-1}(\mathfrak{Y}^A(\mathfrak{r}))\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})) &= \mathfrak{Y}^A(\mathfrak{r}). \end{aligned}$$

If we can show that $\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$, we can then conclude that $\mathfrak{Y}(\mathfrak{r}) = f^{-1}(\mathfrak{Y}^A(\mathfrak{r}))\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}), f(\mathfrak{Y}(\mathfrak{r})) = \mathfrak{Y}^A(\mathfrak{r})$, thus completing the proof of (3). Since $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ is a closed subset of $\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$, the assumption that it is a proper subset permits us to find an element $a_{\mathfrak{R}}$ in $A_{\mathfrak{R}}$ such that $\mathfrak{E}(a_{\mathfrak{R}})$ contains $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ but not $\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$. The closed set $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a_{\mathfrak{R}})$ is then contained in $\mathfrak{E}(\mathfrak{a}_{\mathfrak{R}})$. Hence there exists an element $b_{\mathfrak{R}}$ in $\mathfrak{a}_{\mathfrak{R}}$ such that $\mathfrak{E}(b_{\mathfrak{R}})$ contains $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a_{\mathfrak{R}})$. The element $a_{\mathfrak{R}} \vee b_{\mathfrak{R}}$ then has the property that $\mathfrak{E}(a_{\mathfrak{R}} \vee b_{\mathfrak{R}}) = \mathfrak{E}(a_{\mathfrak{R}}) \vee \mathfrak{E}(b_{\mathfrak{R}})$ contains $\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ but not $\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$. Since $\mathfrak{r}$ is interior to $a_{\mathfrak{R}} \vee b_{\mathfrak{R}}$ by virtue of Theorem 28, the basic ring A contains an element $c = c_{\mathfrak{R}}$ with the property that $\mathfrak{r} \epsilon c'^{-'} = c_{\mathfrak{R}}'^{-'} < (a_{\mathfrak{R}} \vee b_{\mathfrak{R}})'^{-'}$. We therefore have $\mathfrak{X}^A(\mathfrak{r}) \subset \mathfrak{E}(c)$, $\mathfrak{Z}^A(\mathfrak{r}) = f^{-1}(\mathfrak{X}^A(\mathfrak{r})) \subset f^{-1}(\mathfrak{E}(c)) = \mathfrak{E}(c_{\mathfrak{R}})$. Now the relations $(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})'^{-'} \equiv a_{\mathfrak{R}} \vee b_{\mathfrak{R}} (\text{mod } \mathfrak{a}_{\mathfrak{R}})$, $c_{\mathfrak{R}}'^{-'} \equiv c_{\mathfrak{R}} (\text{mod } \mathfrak{a}_{\mathfrak{R}})$, and $c_{\mathfrak{R}}'^{-'} < (a_{\mathfrak{R}} \vee b_{\mathfrak{R}})'^{-'}$ imply the equivalent relations $(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})c_{\mathfrak{R}} \equiv c_{\mathfrak{R}}$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})'c_{\mathfrak{R}} \equiv 0$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $\mathfrak{E}'(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})\mathfrak{E}(c_{\mathfrak{R}}) \subset \mathfrak{E}(\mathfrak{a}_{\mathfrak{R}})$, and $\mathfrak{E}(c_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{E}(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$. Thus we obtain the contradiction

$$\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{E}(c_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{E}(a_{\mathfrak{R}} \vee b_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{E}(a_{\mathfrak{R}} \vee b_{\mathfrak{R}}).$$

We conclude therefore that $\mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) = \mathfrak{X}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$, as we wished to do.

Since no set $\mathfrak{X}^A(\mathfrak{r})$ is contained in $\mathfrak{E}(\mathfrak{a})$, as shown in Theorem 28, the closed sets $\mathfrak{X}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}) = \mathfrak{Y}^A(\mathfrak{r})$ are non-void and constitute a family Υ^A in $\mathfrak{E}'(\mathfrak{a})$ to which Theorem 14 is applicable. By Theorems 3 and 4 we know that $\mathfrak{E}'(\mathfrak{a})$ is a bicompact Boolean subspace of $\mathfrak{E}(A)$, the sets $\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ constituting a basis for it. If $\mathfrak{p}$ is any point of this subspace and if a is an element of A such that $\mathfrak{p} \epsilon \mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$, it is clear from Theorem 28 that there exists a set $\mathfrak{X}^A(\mathfrak{r})$ contained in $\mathfrak{E}(a)$. We therefore have $\mathfrak{Y}^A(\mathfrak{r}) = \mathfrak{X}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}) \subset \mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$. Hence

the family $\mathfrak{Y}^A$ is densely distributed in $\mathfrak{E}'(\mathfrak{a})$ in accordance with Definition 4. If we consider Υ^A as a T_0-space in accordance with Theorem 14, we wish to show that it is topologically independent of A. We therefore compare Υ^A for arbitrary A with the particular space Υ for $A = A_{\mathfrak{R}}$. From (3) above we see that the correspondence $\mathfrak{Y}(\mathfrak{r}) \longleftrightarrow \mathfrak{Y}^A(\mathfrak{r})$ carries Υ into Υ^A in a biunivocal manner, even though we may have $\mathfrak{Y}(\mathfrak{r}) = \mathfrak{Y}(\mathfrak{s})$ or $\mathfrak{Y}^A(\mathfrak{r}) = \mathfrak{Y}^A(\mathfrak{s})$ when $\mathfrak{r} \neq \mathfrak{s}$. We observe next that the subfamilies of Υ specified by the relations $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$, $a_{\mathfrak{R}} = a \epsilon A$, constitute a basis in Υ. If $\mathfrak{G}$ is any relatively open subset of $\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ and if $\mathfrak{Y}(\mathfrak{r})$ is contained in $\mathfrak{G}$, we wish to establish the existence of such an element $a_{\mathfrak{R}}$ with the property that $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{G}$. Since $\mathfrak{X}^A(\mathfrak{r})$ is, by Theorem 26, the intersection of all the sets $\mathfrak{E}(a)$ containing it, the set $\mathfrak{Z}^A(\mathfrak{r}) = f^{-1}(\mathfrak{X}^A(\mathfrak{r}))$ is the intersection of all the corresponding sets $f^{-1}(\mathfrak{E}(a)) = \mathfrak{E}(a_{\mathfrak{R}})$; and $\mathfrak{Y}(\mathfrak{r}) = \mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ is the intersection of all the corresponding sets $\mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$. Let us denote by $\mathfrak{F}$ the closed set which is the complement of $\mathfrak{G}$ in $\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$, and by $\mathfrak{b}$ the class of all elements $a_{\mathfrak{R}} = a \epsilon A$ such that $\mathfrak{X}^A(\mathfrak{r}) \subset \mathfrak{E}(a)$. If we had $\mathfrak{E}(a_{\mathfrak{R}}^{(1)}) \cdots \mathfrak{E}(a_{\mathfrak{R}}^{(n)})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})\mathfrak{F} \neq 0$ for every choice of $a_{\mathfrak{R}}^{(1)}, \cdots, a_{\mathfrak{R}}^{(n)}$ in $\mathfrak{b}$, the bicompactness of $\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ would permit us to write

$$\mathfrak{Y}(\mathfrak{r})\mathfrak{F} = \mathfrak{Z}^A(\mathfrak{r})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})\mathfrak{F} = \prod_{a_{\mathfrak{R}} \epsilon \mathfrak{b}} \mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})\mathfrak{F} \neq 0,$$

contrary to hypothesis. Hence we have $\mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})\mathfrak{F} = \mathfrak{E}(a_{\mathfrak{R}}^{(1)}) \cdots \mathfrak{E}(a_{\mathfrak{R}}^{(n)})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})\mathfrak{F} = 0$ for $a_{\mathfrak{R}} = a_{\mathfrak{R}}^{(1)} \cdots a_{\mathfrak{R}}^{(n)}$ and appropriate elements $a_{\mathfrak{R}}^{(1)}, \cdots, a_{\mathfrak{R}}^{(n)}$ in $\mathfrak{b}$. Since we have $\mathfrak{E}(a) = \mathfrak{E}(a^{(1)}) \cdots \mathfrak{E}(a^{(n)}) \supset \mathfrak{X}^A(\mathfrak{r})$ where $a = a^{(1)} \cdots a^{(n)}$ and $a^{(1)}, \cdots, a^{(n)}$ are the elements $a_{\mathfrak{R}}^{(1)}, \cdots, a_{\mathfrak{R}}^{(n)}$, considered as elements of A, we see that $a_{\mathfrak{R}} = a$ is a member of $\mathfrak{b}$. The desired relation $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{G}$ is thereby established. Since the relations $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(a_{\mathfrak{R}})\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ and $\mathfrak{Y}^A(\mathfrak{r}) \subset \mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ are equivalent by (3) above when $a_{\mathfrak{R}} = a \epsilon A$, we see that the basis just found for Υ is carried by the biunivocal correspondence $\mathfrak{Y}(\mathfrak{r}) \longleftrightarrow \mathfrak{Y}^A(\mathfrak{r})$ into a class of subfamilies in Υ^A which is known from Theorem 23 to be a basis for Υ^A. Hence the spaces Υ and Υ^A are topologically equivalent; and Υ^A is topologically independent of A.

By reference to Theorem 18, we now see that the correspondence $\mathfrak{X}(\mathfrak{r}) \rightarrow \mathfrak{Y}(\mathfrak{r})$ from X to Υ is a univocal continuous correspondence. Thus the spaces $\mathfrak{R}$, $\mathfrak{R}^*$ equivalent to X and Υ respectively have the property that $\mathfrak{R}^*$ is a continuous image of $\mathfrak{R}$. By reference to Theorem 28, we see that, when $\mathfrak{R}$ is an H-space and $\mathfrak{r}$ and $\mathfrak{s}$ are distinct, $\mathfrak{Y}(\mathfrak{r})$ and $\mathfrak{Y}(\mathfrak{s})$ are disjoint. Hence Υ and $\mathfrak{R}^*$ are H-spaces in accordance with Theorem 23; and the correspondences between X and Υ, $\mathfrak{R}$ and $\mathfrak{R}^*$ are biunivocal. To determine the relation of the character of $\mathfrak{R}^*$ to that of $\mathfrak{R}$, we proceed as follows. If $\mathfrak{R}$ has infinite character $\mathfrak{c}$, then there exists a basis for $\mathfrak{R}$ with cardinal number $\mathfrak{c}$.

The Boolean ring A with unit $\mathfrak{R}$ generated by this basis is a basic ring for $\mathfrak{R}$ with cardinal number $\mathfrak{c}$. The Boolean space $\mathfrak{C}(A)$ then has character $\mathfrak{c}$ by virtue of Theorem 1. The character of the subspace $\mathfrak{C}'(\mathfrak{a})$ therefore does not exceed that of $\mathfrak{C}(A)$ or that of $\mathfrak{R}$. The space Υ^A has character not exceeding that of $\mathfrak{C}'(\mathfrak{a})$, as shown in Theorem 23. Since $\mathfrak{R}^*$ is equivalent to Υ^A, we conclude that its character does not exceed the character $\mathfrak{c}$ of $\mathfrak{R}$. Finally the suppression of the open set $\mathfrak{C}(\mathfrak{a})$ from the map $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X}^A)$ yields the map $m(\mathfrak{R}^*, \mathfrak{C}'(\mathfrak{a}), \Upsilon^A)$ in accordance with Definition 7; and the latter map is irredundant in accordance with Definition 9 and Theorem 20.

The preceding theorem raises several questions which we shall state and consider later, in Chapter III.

3. **Relations between algebraic and other maps.** In various applications of the theory of maps, it is essential to have information about the relations of general Boolean maps to the algebraic Boolean maps of the preceding section. The analysis of such relations appears to be quite difficult. In any case we have not succeeded in outlining a comprehensive survey of the possible connections. We shall therefore confine ourselves in the present section to the consideration of a few special results which find immediate application in the sequel. These results are concerned with a new concept in the theory of maps, introduced now by the following definition.

DEFINITION 12. *If $\mathfrak{X}$ is any non-void family of distinct non-void closed sets $\mathfrak{x}$ in a T_1-space $\mathfrak{S}$, then a set $\mathfrak{F}$ in $\mathfrak{S}$ is said to be an $\mathfrak{X}$-set when*

(1) *$\mathfrak{F}$ is closed and non-void*; (2) *every open set $\mathfrak{G}$ in $\mathfrak{S}$ which contains $\mathfrak{F}$ also contains some member of the family $\mathfrak{X}$.*

An $\mathfrak{X}$-set is said to be minimal (*with respect to the property P*) *if it is an $\mathfrak{X}$-set* (*with property P*) *and contains no proper $\mathfrak{X}$-subset* (*with property P*).

It is obvious that every member of the family $\mathfrak{X}$ is an $\mathfrak{X}$-set; but the chief interest of the definition lies in the possibility that there exist $\mathfrak{X}$-sets not belonging to $\mathfrak{X}$. We may begin by considering the determination of $\mathfrak{X}$-sets when $\mathfrak{S}$ is a bicompact Boolean space.

THEOREM 30. *If the space $\mathfrak{S}$ of Definition 12 is a bicompact Boolean space $\mathfrak{C}(A)$ representing a Boolean ring A with unit, the open sets $\mathfrak{G}$ of that definition may be restricted to be bicompact without modifying the content of the definition. If $\mathfrak{c}$ is the class of all elements a in A such that $\mathfrak{C}(a)$ contains a given $\mathfrak{X}$-set $\mathfrak{F}$, then $\mathfrak{c}$ has the properties*

(1) *if $a_1, \cdots, a_n$ are in $\mathfrak{c}$, then $a_1 \cdots a_n$ is in $\mathfrak{c}$;*

(2) *if $a_1, \cdots, a_n$ are in $\mathfrak{c}$, then $\mathfrak{C}(a)$, where $a = a_1 \cdots a_n$, contains some member of $\mathfrak{X}$;*

and the relation $\mathfrak{F} = \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a)$ is valid. Conversely, if $\mathfrak{c}$ is any non-void subclass of A with the properties (1) *and* (2), *then the set $\mathfrak{F} = \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a)$ is an $\mathfrak{X}$-set.*

If the open set $\mathfrak{G}$ contains the closed set $\mathfrak{F}$ in $\mathfrak{S}=\mathfrak{C}(A)$, then there exists an element a in A such that $\mathfrak{F}\subset\mathfrak{C}(a)\subset\mathfrak{G}$. In fact, if $\mathfrak{s}$ is any point of $\mathfrak{F}$, there exists an element $a(\mathfrak{s})$ in A such that $\mathfrak{s}\epsilon\mathfrak{C}(a(\mathfrak{s}))\subset\mathfrak{G}$; and the bicompactness of $\mathfrak{F}$ permits us to take $a=a(\mathfrak{s}_1)\vee\cdots\vee a(\mathfrak{s}_n)$, $\mathfrak{C}(a)=\mathfrak{C}(a(\mathfrak{s}_1))\cup\cdots\cup\mathfrak{C}(a(\mathfrak{s}_n))$ for a suitable choice of points $\mathfrak{s}_1,\cdots,\mathfrak{s}_n$ in $\mathfrak{F}$. Hence the restriction imposed by requiring that $\mathfrak{G}=\mathfrak{C}(a)$ in Definition 12 does not modify the content of the definition. If $\mathfrak{c}$ is the class of all elements a such that $\mathfrak{C}(a)\supset\mathfrak{F}$, then the properties (1) and (2) are easily established as follows: if $a=a_1\cdots a_n$ where $a_1,\cdots,a_n$ are in $\mathfrak{c}$, then $\mathfrak{C}(a)=\mathfrak{C}(a_1)\cdots\mathfrak{C}(a_n)\supset\mathfrak{F}$; thus a is in $\mathfrak{c}$ and $\mathfrak{C}(a)$ contains a member of $\mathfrak{X}$ since it contains the $\mathfrak{X}$-set $\mathfrak{F}$. If $\mathfrak{s}$ is any point which does not belong to $\mathfrak{F}$, the open set $\mathfrak{G}=\mathfrak{C}(A)-\{\mathfrak{s}\}$ contains $\mathfrak{F}$; and there exists an element a, necessarily in $\mathfrak{c}$, such that $\mathfrak{F}\subset\mathfrak{C}(a)\subset\mathfrak{C}(A)-\{\mathfrak{s}\}$. Hence we see that $\mathfrak{F}=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$. We pass now to the converse. The properties (1) and (2) assumed for the class $\mathfrak{c}$ show immediately that when $a_1,\cdots,a_n$ are in $\mathfrak{c}$ the set $\mathfrak{C}(a_1)\cdots\mathfrak{C}(a_n)=\mathfrak{C}(a_1\cdots a_n)$ is non-void. Since $\mathfrak{C}(A)$ is bicompact and the sets $\mathfrak{C}(a)$ are closed, the intersection $\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$ is a non-void closed set $\mathfrak{F}$. If $\mathfrak{C}(b)\supset\mathfrak{F}$, where b is an arbitrary element in A, we must have $\mathfrak{C}(a_1)\cdots\mathfrak{C}(a_n)\mathfrak{C}'(b)=0$ for some elements $a_1,\cdots,a_n$ in $\mathfrak{c}$: for otherwise we would have

$$0=\mathfrak{F}\mathfrak{C}'(b)=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)\mathfrak{C}'(b)\neq 0.$$

Hence there exists an element $a=a_1\cdots a_n$ in $\mathfrak{c}$ such that $\mathfrak{F}\subset\mathfrak{C}(a)\subset\mathfrak{C}(b)$. Property (2) shows that $\mathfrak{C}(a)$, and hence also $\mathfrak{C}(b)$, contains some member of $\mathfrak{X}$. Since b was subject only to the restriction $\mathfrak{C}(b)\supset\mathfrak{F}$, it follows that $\mathfrak{F}$ is an $\mathfrak{X}$-set in accordance with Definition 12 and the first part of the present theorem.

We next consider the existence of minimal $\mathfrak{X}$-sets.

THEOREM 31. *Let P be a property such that the members of any descending transfinite sequence of sets with the property P contain a common subset with the property P. Then every set with the property P contains a minimal set with the property P.*

Let Ω be the first ordinal number such that the class of ordinals $\alpha<\Omega$ has cardinal number exceeding that of the class of all sets with the property P. We then define a "minimizing" sequence $\mathfrak{F}_\alpha$, $\alpha<\Omega$, of sets with the property P such that: (1) $\alpha>\beta$ implies $\mathfrak{F}_\alpha\subset\mathfrak{F}_\beta$; (2) if $\prod_{\beta<\alpha}\mathfrak{F}_\beta$ contains a set which has the property P but which is not minimal, then $\mathfrak{F}_\alpha$ is a proper subset of one such set. We choose $\mathfrak{F}_1$ as an arbitrary set with the property P. Then if $\mathfrak{F}_\beta$ has been defined for $\beta<\alpha<\Omega$, we form the intersection $\prod_{\beta<\alpha}\mathfrak{F}_\beta$. By hypothesis there exists a set $\mathfrak{F}$ which has the property P and is contained in this

intersection. Now if every such $\mathfrak{F}$ is minimal, we put $\mathfrak{F}_\alpha=\mathfrak{F}$; and otherwise we choose $\mathfrak{F}_\alpha$ as a proper subset of some $\mathfrak{F}$ so that it has the property P. By the principle of transfinite induction $\mathfrak{F}_\alpha$ is thereby defined for $\alpha<\Omega$. Now the class of ordinals α such that $\mathfrak{F}_\alpha=\mathfrak{F}_\gamma$ for some ordinal γ in the range $\alpha<\gamma<\Omega$ is a non-void class by virtue of our choice of Ω. This class therefore has a first member β. The relations $\beta<\beta+1\leqq\gamma$, $\mathfrak{F}_\beta=\mathfrak{F}_\gamma$, imply $\mathfrak{F}_\beta\supset\mathfrak{F}_{\beta+1}\supset\mathfrak{F}_\gamma\supset\mathfrak{F}_\beta$, $\mathfrak{F}_\beta=\mathfrak{F}_{\beta+1}$. Since

$$\prod_{\alpha<\beta+1}\mathfrak{F}_\alpha=\prod_{\alpha\leqq\beta}\mathfrak{F}_\alpha=\mathfrak{F}_\beta,$$

the relation $\mathfrak{F}_\beta=\mathfrak{F}_{\beta+1}$ implies in accordance with (2) that $\mathfrak{F}_{\beta+1}$ is minimal. Since it is evident that $\mathfrak{F}_{\beta+1}\subset\mathfrak{F}_1$, the theorem is established.

THEOREM 32. *If the space $\mathfrak{S}$ of Definition* 12 *is a bicompact H-space, then each of the following properties*:

(1) *the property of being an $\mathfrak{X}$-set*;

(2) *the property of being an $\mathfrak{X}$-set containing a given point $\mathfrak{s}$*;

(3) *the property of being an $\mathfrak{X}$-set contained in a given set*;

is a property P of the kind described in the preceding theorem. Hence minimal sets with respect to each of them exist in accordance with that theorem.

Let P be any one of the three properties. Then it is sufficient for us to show that if $\mathfrak{F}_\alpha$, $\alpha<\omega$, is a transfinite sequence such that (1) $\mathfrak{F}_\alpha$ has the property P, and (2) $\mathfrak{F}_\alpha\subset\mathfrak{F}_\beta$ when $\alpha>\beta$, then $\prod_{\alpha<\omega}\mathfrak{F}_\alpha$ also has the property P. Since $\mathfrak{F}_\alpha$ is an $\mathfrak{X}$-set, it is closed and non-void. Hence, by the assumption concerning $\mathfrak{S}$, the intersection $\prod_{\alpha<\omega}\mathfrak{F}_\alpha$ is also closed and non-void. In order to show that this set has the required property, it is enough to prove that it is an $\mathfrak{X}$-set. If $\mathfrak{G}$ is any open set containing it, we have

$$\prod_{\alpha<\omega}\mathfrak{F}_\alpha\mathfrak{G}'=\mathfrak{G}'\prod_{\alpha<\omega}\mathfrak{F}_\alpha=0$$

where the sets $\mathfrak{F}_\alpha\mathfrak{G}'$ are closed and satisfy the relation $\mathfrak{F}_\alpha\mathfrak{G}'\subset\mathfrak{F}_\beta\mathfrak{G}'$ when $\alpha>\beta$. Our assumption concerning $\mathfrak{S}$ now implies that there exists a first ordinal number α preceding ω for which $\mathfrak{F}_\alpha\mathfrak{G}'=0$ or, equivalently, $\mathfrak{F}_\alpha\subset\mathfrak{G}$. The fact that $\mathfrak{F}_\alpha$ is an $\mathfrak{X}$-set shows that $\mathfrak{G}$ contains some member of $\mathfrak{X}$. Since the open set $\mathfrak{G}$ was subject only to the restriction that it contain $\prod_{\alpha<\omega}\mathfrak{F}_\alpha$, the latter set is an $\mathfrak{X}$-set in accordance with Definition 12. This completes the proof.

We may give without proof the following useful result.

THEOREM 33. *If the families $\mathfrak{X}$ and $\mathfrak{Y}$ in the space $\mathfrak{S}$ satisfy the relation $\mathfrak{X}\supset\mathfrak{Y}$, then every $\mathfrak{Y}$-set is an $\mathfrak{X}$-set; and, if $\mathfrak{S}$ is a bicompact H-space, every*

minimal Υ-set contains a minimal $\mathfrak{X}$-set, every Υ-set minimal with respect to the property of containing a given point $\mathfrak{s}$ contains an $\mathfrak{X}$-set minimal with respect to the same property.

Naturally, it is of considerable interest to determine the $\mathfrak{X}$-sets and minimal $\mathfrak{X}$-sets in the case of the family $\mathfrak{X}$ in an algebraic map. We find the following situation.

THEOREM 34. *If $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ is an algebraic Boolean map, then the $\mathfrak{X}$-sets are characterized as the closed subsets of $\mathfrak{B}=\mathfrak{C}(A)$ which have at least one point in common with the set $\mathfrak{C}'(\mathfrak{a})$, where $\mathfrak{a}$ is the ideal of nowhere dense sets in the basic ring A. The minimal $\mathfrak{X}$-sets are characterized as the one-element subsets of $\mathfrak{C}'(\mathfrak{a})$. The $\mathfrak{X}$-sets minimal with respect to the property of containing a given point $\mathfrak{p}$ in $\mathfrak{C}(\mathfrak{a}) \subset \mathfrak{B}=\mathfrak{C}(A)$ are characterized as the sets containing two points, the point $\mathfrak{p}$ and a point of $\mathfrak{C}'(\mathfrak{a})$.*

First, let us show that every $\mathfrak{X}$-set $\mathfrak{F}$ has at least one point in common with $\mathfrak{C}'(\mathfrak{a})$. By Theorem 30, we have $\mathfrak{F}=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$ where $\mathfrak{c}$ is the class of all a in A for which $\mathfrak{F}\subset\mathfrak{C}(a)$. By virtue of the bicompactness of $\mathfrak{B}=\mathfrak{C}(A)$, we see that

$$\mathfrak{F}\mathfrak{C}'(\mathfrak{a}) = \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a)\mathfrak{C}'(\mathfrak{a}) = 0$$

if and only if $\mathfrak{C}(a)\mathfrak{C}'(\mathfrak{a})=0$ for some a in $\mathfrak{c}$. By Theorem 28, we know that $\mathfrak{C}(a)$ contains a member of $\mathfrak{X}$ if and only if $\mathfrak{C}(a)\mathfrak{C}'(\mathfrak{a})\neq 0$. Since $\mathfrak{F}$ is an $\mathfrak{X}$-set, the relation $\mathfrak{F}\subset\mathfrak{C}(a)$ implies that $\mathfrak{C}(a)$ contains some member of $\mathfrak{X}$. Hence we conclude that $\mathfrak{F}\mathfrak{C}'(\mathfrak{a})\neq 0$, as we wished to prove. If $\mathfrak{p}$ is any point of $\mathfrak{C}'(\mathfrak{a})$, then $\mathfrak{p}\epsilon\mathfrak{C}(a)$ implies that $\mathfrak{C}(a)$ contains some member of $\mathfrak{X}$; and the one-element set $\{\mathfrak{p}\}$ is therefore an $\mathfrak{X}$-set. From these facts, the remaining statements of the theorem can be deduced in an obvious way.

As a consequence of this theorem, we have the following comment on Theorem 29.

THEOREM 35. *In the notation of Theorem 29, every Z^A-set in $\mathfrak{C}(A_{\mathfrak{R}})$ is an $\mathfrak{X}$-set.*

If $\mathfrak{F}$ is a Z^A-set and $\mathfrak{C}(a)\supset\mathfrak{F}$ where $a\epsilon A_{\mathfrak{R}}$, then $\mathfrak{C}(a)$ contains some set $\mathfrak{Z}^A(\mathfrak{r})$. Since $\mathfrak{Z}^A(\mathfrak{r})\supset\mathfrak{X}(\mathfrak{r})$, it follows that $\mathfrak{F}$ is an $\mathfrak{X}$-set.

We now proceed to study a natural method for relating a general Boolean map to an algebraic map.

THEOREM 36. *Let $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ be an arbitrary Boolean map where the bicompact Boolean space $\mathfrak{B}$ is the representative of a Boolean ring A. Let $a^*=\mathfrak{G}(a)$ be the open subset of $\mathfrak{R}$ determined from the given map through the relation*

$\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(a)$, where a is an arbitrary element of A. Let A^* be the basic ring generated by the basis of all sets $a^* = \mathfrak{G}(a)$ in $\mathfrak{R}$; and let $m(\mathfrak{R}, \mathfrak{B}^*, \mathcal{X}^*)$ be the algebraic map defined by A^* in $\mathfrak{B}^* = \mathfrak{E}(A^*)$. Let $\mathfrak{F}$ be the class of all $\mathcal{X}$-sets in $\mathfrak{B}$, $\mathfrak{F}^*$ the class of all $\mathcal{X}^*$-sets in $\mathfrak{B}^*$. If $\mathfrak{F}$ is an arbitrary $\mathcal{X}$-set and $\mathfrak{c}$ is the class of all a in A such that $\mathfrak{E}(a) \supset \mathfrak{F}$, then the set $f(\mathfrak{F}) = \prod_{a \epsilon \mathfrak{c}} \mathfrak{E}(a^*)$ is an $\mathcal{X}^*$-set. The correspondence $\mathfrak{F} \to f(\mathfrak{F})$ has the properties:

(1) *f carries $\mathfrak{F}$ univocally into a subclass $\mathfrak{F}^{**}$ of $\mathfrak{F}^*$;*

(2) *$\mathfrak{F}_1 \supset \mathfrak{F}_2$ implies $f(\mathfrak{F}_1) \supset f(\mathfrak{F}_2)$;*

(3) *the intersection of $\mathcal{X}$-sets with the property $f(\mathfrak{F}) = \mathfrak{F}^*$, where $\mathfrak{F}^*$ is a fixed member of $\mathfrak{F}^{**}$, is an $\mathcal{X}$-set with this property; and the intersection of all sets with the property is the unique $\mathcal{X}$-set minimal with respect to this property;*

(4) *$\mathfrak{X}(\mathfrak{r})$ is the minimal $\mathcal{X}$-set with the property $f(\mathfrak{F}) = \mathfrak{X}^*(\mathfrak{r})$;*

(5) *f carries the minimal sets in $\mathfrak{F}$ biunivocally into the minimal sets in $\mathfrak{F}^{**}$.*

If the class $\mathfrak{F}^{**}$ contains the minimal sets in $\mathfrak{F}^*$ or, more generally, if each point of the set $\mathfrak{E}'(\mathfrak{a}^*)$, where $\mathfrak{a}^*$ is the ideal of nowhere dense sets in A^*, belongs to exactly one minimal set in $\mathfrak{F}^{**}$, the map $m(\mathfrak{R}, \mathfrak{B}, \mathcal{X})$ has the following properties:

(1) *the union of all minimal $\mathcal{X}$-sets is an $\mathcal{X}$-set $\mathfrak{F}(\mathcal{X})$;*

(2) *every $\mathcal{X}$-set has at least one point in common with the set $\mathfrak{F}(\mathcal{X})$;*

(3) *if $\mathfrak{E}(a) \subset \mathfrak{F}'(\mathcal{X})$, then $\mathfrak{E}(a)$ contains no member of $\mathcal{X}$.*

Since the set $\mathfrak{F}(\mathcal{X})$ is closed, it is bicompact.

The preliminary justification of the constructions leading to the algebraic map described here can be obtained by reference to Theorems 14, 23, and 25 and to Definitions 7, 10, and 11. We begin our proof by a consideration of the definition of the correspondence f. By Theorem 30, we have $\mathfrak{F} = \prod_{a \epsilon \mathfrak{c}} \mathfrak{E}(a)$, the notations being those introduced above. The elements $a^* = \mathfrak{G}(a)$ where $a \epsilon \mathfrak{c}$ constitute a class $\mathfrak{c}^*$ in A^* which has properties (1) and (2) of Theorem 30. For $a^* = a_1^* \cdots a_n^* = \mathfrak{G}(a_1) \cdots \mathfrak{G}(a_n) = \mathfrak{G}(a_1 \cdots a_n) = \mathfrak{G}(a)$ where $a = a_1 \cdots a_n$ is in $\mathfrak{c}$; and $a^* = \mathfrak{G}(a)$, $a \epsilon \mathfrak{c}$, implies the existence of a point $\mathfrak{r}$ in $\mathfrak{R}$ satisfying the equivalent relations $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(a)$, $\mathfrak{r} \epsilon \mathfrak{G}(a)$, $\mathfrak{r} \epsilon (a^*)'^{-\prime} = a^*$, $\mathfrak{X}^*(\mathfrak{r}) \subset \mathfrak{E}(a^*)$. Hence Theorem 30 shows that $f(\mathfrak{F}) = \prod_{a \epsilon \mathfrak{c}} \mathfrak{E}(a^*)$ is an $\mathcal{X}^*$-set. Properties (1) and (2) of the correspondence f are obvious from the definition. Next we consider property (3). If A is an abstract class of elements α to each of which corresponds a member $\mathfrak{F}_\alpha$ of $\mathfrak{F}$ such that $f(\mathfrak{F}_\alpha) = \mathfrak{F}^* \epsilon \mathfrak{F}^{**}$, we wish first to prove that $\mathfrak{F} = \prod_{\alpha \epsilon A} \mathfrak{F}_\alpha$ is an $\mathcal{X}$-set with the property $f(\mathfrak{F}) = \mathfrak{F}^*$. We again appeal to Theorem 30. We begin by writing $\mathfrak{F}_\alpha = \prod_{a \epsilon \mathfrak{c}(\alpha)} \mathfrak{E}(a)$, where $\mathfrak{c}(\alpha)$ is the class of all a in A such that $\mathfrak{E}(a) \supset \mathfrak{F}_\alpha$. We then define $\mathfrak{c}$ as the class of all elements $a = a_1 \cdots a_n$ where $a_k \epsilon \mathfrak{c}(\alpha_k)$ for $k = 1, \cdots, n$ and $n = 1, 2, 3, \cdots$. It is then evident that $\mathfrak{c}$ has property (1) of Theorem 30. In addition

we can prove that $\mathfrak{c}$ has property (2). We first take the arbitrary element a in $\mathfrak{c}$ and express it in the form $a=a_1 \cdots a_n$, where $a_k \epsilon \mathfrak{c}(\alpha_k)$, noted above. We then have $a^*=\mathfrak{G}(a)=\mathfrak{G}(a_1)\cdots\mathfrak{G}(a_n)=a_1^*\cdots a_n^*$, and $\mathfrak{C}(a^*)=\mathfrak{C}(a_1^*)\cdots\mathfrak{C}(a_n^*)$. By hypothesis $\mathfrak{C}(a_k^*)\supset f(\mathfrak{F}_{\alpha_k})=\mathfrak{F}^*$ for $k=1,\cdots,n$. Hence $\mathfrak{C}(a^*)$ also contains $\mathfrak{F}^*$. Since $\mathfrak{F}^*$ is an $\mathcal{X}^*$-set, there exists a point $\mathfrak{r}$ in $\mathfrak{R}$ satisfying the equivalent relations $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{C}(a^*)$, $\mathfrak{r}\epsilon(a^*)'^{-\prime}=a^*=\mathfrak{G}(a)$, $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{C}(a)$. Since $\mathfrak{c}$ thus has properties (1) and (2) of Theorem 30, we conclude that

$$\mathfrak{F}=\prod_{\alpha\epsilon A}\mathfrak{F}_\alpha=\prod_{\alpha\epsilon A}\prod_{a\epsilon\mathfrak{c}(\alpha)}\mathfrak{C}(a)=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$$

is an $\mathcal{X}$-set. As we saw above, $a\epsilon\mathfrak{c}$ implies $\mathfrak{C}(a^*)\supset\mathfrak{F}^*$; and it follows that $f(\mathfrak{F})\supset\mathfrak{F}^*$. On the other hand the relation $\mathfrak{F}\subset\mathfrak{F}_\alpha$, $\alpha\epsilon A$, implies $f(\mathfrak{F})\subset f(\mathfrak{F}_\alpha)=\mathfrak{F}^*$. Hence we have $f(\mathfrak{F})=\mathfrak{F}^*$, as we wished to prove. It is now evident that the intersection of all sets $\mathfrak{F}$ such that $f(\mathfrak{F})=\mathfrak{F}^*$ is the unique $\mathcal{X}$-set minimal with respect to this property. Part of the property (5) follows immediately from (2) and (3). If $\mathfrak{F}^*$ is minimal in the family $\mathfrak{f}^{**}$, then the set $\mathfrak{F}$ in $\mathfrak{f}$ which is minimal with respect to the property $f(\mathfrak{F})=\mathfrak{F}^*$ exists by (3); and it must be minimal in the entire family $\mathfrak{f}$, since an $\mathcal{X}$-set $\mathfrak{F}_1$ contained in $\mathfrak{F}$ has the properties $f(\mathfrak{F}_1)\subset f(\mathfrak{F})=\mathfrak{F}^*$, $f(\mathfrak{F}_1)=\mathfrak{F}^*$, $\mathfrak{F}_1\supset\mathfrak{F}$, and therefore coincides with $\mathfrak{F}$. The rest of property (4) is established by reasoning similar to that applied to prove (3). If $\mathfrak{F}$ is minimal in $\mathfrak{f}$, and $\mathfrak{F}_1^*$ is a member of $\mathfrak{f}^{**}$ contained in $\mathfrak{F}^*=f(\mathfrak{F})$, then there exists a set $\mathfrak{F}_1$ in $\mathfrak{f}$ such that $f(\mathfrak{F}_1)=\mathfrak{F}_1^*$. We let $\mathfrak{c}$ be the class of all elements a in A such that $a=bb_1$ where $\mathfrak{C}(b)\supset\mathfrak{F}$, $\mathfrak{C}(b_1)\supset\mathfrak{F}_1$. Then $\mathfrak{F}\mathfrak{F}_1=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$. If $a=bb_1$ is any element in $\mathfrak{c}$, we have $a^*=b^*b_1^*$, $\mathfrak{C}(a^*)=\mathfrak{C}(b^*)\mathfrak{C}(b_1^*)$. By hypothesis $\mathfrak{C}(a^*)\supset\mathfrak{F}_1^*$. Since $\mathfrak{F}_1^*$ is an $\mathcal{X}^*$-set, there exists a point $\mathfrak{r}$ in $\mathfrak{R}$ satisfying the equivalent relations $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{C}(a^*)$, $\mathfrak{r}\epsilon(a^*)'^{-\prime}=a^*=\mathfrak{G}(a)$, $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{C}(a)$. Since $\mathfrak{c}$ thus has properties (1) and (2) of Theorem 30, we conclude that $\mathfrak{F}\mathfrak{F}_1$ is an $\mathcal{X}$-set. By virtue of the fact that $\mathfrak{F}$ is a minimal $\mathcal{X}$-set, we must have $\mathfrak{F}\mathfrak{F}_1=\mathfrak{F}$, $\mathfrak{F}_1\supset\mathfrak{F}$, $f(\mathfrak{F}_1)\supset f(\mathfrak{F})$, $\mathfrak{F}_1^*\supset\mathfrak{F}^*$, and $\mathfrak{F}_1^*=\mathfrak{F}^*$. Thus $\mathfrak{F}^*=f(\mathfrak{F})$ is minimal in $\mathfrak{f}^{**}$, as we wished to prove. If $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are minimal $\mathcal{X}$-sets such that $f(\mathfrak{F}_1)=f(\mathfrak{F}_2)$, then by (3) $\mathfrak{F}_1\mathfrak{F}_2$ is an $\mathcal{X}$-set. Consequently $\mathfrak{F}_1\mathfrak{F}_2=\mathfrak{F}_1$, $\mathfrak{F}_1\mathfrak{F}_2=\mathfrak{F}_2$ and $\mathfrak{F}_1=\mathfrak{F}_2$. Thus f defines a biunivocal correspondence between the minimal sets in $\mathfrak{f}$ and those in $\mathfrak{f}^{**}$. This completes the proof of property (5). To prove property (4) we use Theorem 30 once again. We have $\mathfrak{X}(\mathfrak{r})=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a)$, where $\mathfrak{c}$ is the class of all a in A such that $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{C}(a)$. Since $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{C}(a)$, $\mathfrak{r}\epsilon\mathfrak{G}(a)=a^*=(a^*)'^{-\prime}$, and $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{C}(a^*)$ are equivalent relations, we see that

$$f(\mathfrak{X}(\mathfrak{r}))=\prod_{a\epsilon\mathfrak{c}}\mathfrak{C}(a^*)=\mathfrak{X}^*(\mathfrak{r}),$$

with appropriate reference to Theorem 26. On the other hand, if $\mathfrak{F}$ is any $\mathfrak{X}$-set with the property $f(\mathfrak{F}) = \mathfrak{X}^*(\mathfrak{r})$, we express $\mathfrak{F}$ as an intersection of sets $\mathfrak{C}(a)$ containing $\mathfrak{F}$. We then have $a^* = \mathfrak{G}(a)$, $\mathfrak{C}(a^*) \supset \mathfrak{X}^*(\mathfrak{r})$ for such elements a, in accordance with the definition of the correspondence f. The second of these relations implies $\mathfrak{r}\epsilon(a^*)'^{-\prime} = a^* = \mathfrak{G}(a)$, $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a)$. Hence $\mathfrak{F}$ contains $\mathfrak{X}(\mathfrak{r})$; and $\mathfrak{X}(\mathfrak{r})$ is the unique $\mathfrak{X}$-set minimal with respect to the property $f(\mathfrak{F}) = \mathfrak{X}^*(\mathfrak{r})$.

Finally, we come to the special property described in the last part of the theorem. We begin by considering the class $\mathfrak{c}$ of all elements a in A such that $\mathfrak{C}(a^*)$, where $a^* = \mathfrak{G}(a)$, contains a given point $\mathfrak{p}$ in $\mathfrak{C}'(\mathfrak{a}^*)$. If $a_1, \cdots, a_n$ are in $\mathfrak{c}$, then the element $a = a_1 \cdots a_n$ has the properties $a^* = \mathfrak{G}(a) = \mathfrak{G}(a_1) \cdots \mathfrak{G}(a_n) = a_1^* \cdots a_n^*$, $\mathfrak{C}(a^*) = \mathfrak{C}(a_1^*) \cdots \mathfrak{C}(a_n^*)$, $\mathfrak{p}\epsilon\mathfrak{C}(a^*)$, and therefore belongs to $\mathfrak{c}$. If a is any element in $\mathfrak{c}$, then the relation $\mathfrak{p}\epsilon\mathfrak{C}(a^*)$ implies the existence of a point $\mathfrak{r}$ in $\mathfrak{R}$ such that $\mathfrak{X}^*(\mathfrak{r}) \subset \mathfrak{C}(a^*)$, by virtue of the fact that $\{\mathfrak{p}\}$ is an $\mathfrak{X}^*$-set in accordance with Theorem 34. Since $\mathfrak{X}^*(\mathfrak{r}) \subset \mathfrak{C}(a^*)$ implies $\mathfrak{r}\epsilon(a^*)'^{-\prime} = a^* = \mathfrak{G}(a)$ and $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a)$, we conclude that $\mathfrak{c}$ has properties (1) and (2) of Theorem 30. Hence the set $\mathfrak{F} = \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a)$ is an $\mathfrak{X}$-set, and its correspondent $\mathfrak{F}^* = f(\mathfrak{F}) = \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a^*)$ contains the point $\mathfrak{p}$. If $\mathfrak{F}_1$ is any $\mathfrak{X}$-set such that $\mathfrak{p}\epsilon f(\mathfrak{F}_1) = \mathfrak{F}_1^*$, we have

$$\mathfrak{F}_1 = \prod_{a\epsilon\mathfrak{c}_1} \mathfrak{C}(a), \qquad \mathfrak{F}_1^* = \prod_{a\epsilon\mathfrak{c}_1} \mathfrak{C}(a^*) \supset \prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a^*) = \mathfrak{F}^*,$$

where $\mathfrak{c}_1$ is the class of all a in A such that $\mathfrak{F}_1 \subset \mathfrak{C}(a)$. Hence we see that the set $\mathfrak{F}^*$ is the unique set in $\mathfrak{F}^{**}$ minimal with respect to the property of containing the given point $\mathfrak{p}$. Thus if a minimal set in $\mathfrak{F}^{**}$ contains the point $\mathfrak{p}$, it must coincide with $\mathfrak{F}^*$. We now assume that each point $\mathfrak{p}$ in $\mathfrak{C}'(\mathfrak{a}^*)$ belongs to exactly one minimal set in $\mathfrak{F}^{**}$. Let us consider the union $\mathfrak{F}(\mathfrak{X})$ of all minimal sets in $\mathfrak{F}$. If $\mathfrak{s}$ is any point in $\mathfrak{F}'(\mathfrak{X})$, a familiar argument shows that every minimal $\mathfrak{X}$-set $\mathfrak{F}$ determines an element a in A such that $\mathfrak{F} \subset \mathfrak{C}(a)$, $\mathfrak{s}\epsilon\mathfrak{C}'(a)$. The corresponding element $a^* = \mathfrak{G}(a)$ then has the property $\mathfrak{C}(a^*) \supset f(\mathfrak{F}) = \mathfrak{F}^*$. By property (5) above, $\mathfrak{F}^*$ is a minimal set in $\mathfrak{F}^{**}$. By our assumption concerning the minimal sets in $\mathfrak{F}^{**}$, and by further reference to property (5), we see that the sets $\mathfrak{C}(a^*)$ thus obtained cover $\mathfrak{C}'(\mathfrak{a}^*)$. Because $\mathfrak{C}'(\mathfrak{a}^*)$ is bicompact, there exist elements $a_1, \cdots, a_n$ such that the corresponding sets $\mathfrak{C}(a_1^*), \cdots, \mathfrak{C}(a_n^*)$ cover $\mathfrak{C}'(\mathfrak{a}^*)$. If $\mathfrak{F}^*$ is any minimal set in $\mathfrak{F}^{**}$, it contains a point $\mathfrak{p}$ in $\mathfrak{C}'(\mathfrak{a}^*)$; and this point belongs to no other minimal set in $\mathfrak{F}^{**}$. If $\mathfrak{p}\epsilon\mathfrak{C}(a_k^*)$, a relation which must be satisfied for some index k, $k = 1, \cdots, n$, we see that $\mathfrak{F}^* \subset \mathfrak{C}(a_k^*)$ in accordance with the results proved above. Furthermore, there exists a unique minimal $\mathfrak{X}$-set $\mathfrak{F}$ such that $f(\mathfrak{F}) = \mathfrak{F}^*$. As we saw above, it is necessary that $\mathfrak{F} \subset \mathfrak{C}(a_k)$: for the set $\prod_{a\epsilon\mathfrak{c}} \mathfrak{C}(a)$, where $\mathfrak{p}\epsilon\mathfrak{C}(a^*)$ characterizes the class $\mathfrak{c}$, has $\mathfrak{F}^*$ as its correspondent

and is contained in $\mathfrak{C}(a_k)$. Now this result shows that every minimal $\mathfrak{X}$-set is contained in some of the sets $\mathfrak{C}(a_k)$, $k=1, \cdots, n$. Hence, if we put $a(\mathfrak{s})=a_1 \vee \cdots \vee a_n$, we have $\mathfrak{F}(\mathfrak{X}) \subset \mathfrak{C}(a(\mathfrak{s}))$, $\mathfrak{s}\epsilon\mathfrak{C}'(a(\mathfrak{s}))$. The set $\mathfrak{F}(\mathfrak{X})$ therefore coincides with the intersection of the closed sets $\mathfrak{C}(a(\mathfrak{s}))$ where $\mathfrak{s}\epsilon\mathfrak{F}'(\mathfrak{X})$; accordingly it must be closed and bicompact. Since $\mathfrak{F}(\mathfrak{X})$ contains $\mathfrak{X}$-sets, namely, the minimal $\mathfrak{X}$-sets, it is obviously itself an $\mathfrak{X}$-set. Since every $\mathfrak{X}$-set contains a minimal $\mathfrak{X}$-set, the remaining properties stated in the theorem are obvious.

In order to apply the preceding theorem, we shall appeal to the following result.

THEOREM 37. *Let $\mathfrak{S}$ be any non-void subspace of a T_0-space $\mathfrak{R}$; let $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ be the complete algebraic map of $\mathfrak{R}$ in $\mathfrak{B}=\mathfrak{C}(A_{\mathfrak{R}})$; and let $\mathfrak{B}(\mathfrak{S})$ be the closure of the union of all sets $\mathfrak{X}(\mathfrak{s})$ where $\mathfrak{s}\epsilon\mathfrak{S}$. Then the Boolean map defined by the family $\mathfrak{X}(\mathfrak{S})$ of all $\mathfrak{X}(\mathfrak{s})$, where $\mathfrak{s}\epsilon\mathfrak{S}$ is a map $m(\mathfrak{S}, \mathfrak{B}(\mathfrak{S}), \mathfrak{X}(\mathfrak{S}))$ with the special property described in Theorem* 36.

If $\mathfrak{G}$ is an arbitrary open subset of $\mathfrak{S}$, there exists an open set a in $A_{\mathfrak{R}}$ such that $a\mathfrak{S}=\mathfrak{G}$. It is then clear that, for $\mathfrak{s}$ in $\mathfrak{S}$, we have $\mathfrak{X}(\mathfrak{s}) \subset \mathfrak{C}(a)$ if and only if $\mathfrak{s}\epsilon\mathfrak{G}$. The set $\mathfrak{C}(a)\mathfrak{B}(\mathfrak{S})$ is a bicompact open subset of $\mathfrak{B}(\mathfrak{S})$ and therefore represents an element b of the Boolean ring A which has $\mathfrak{B}(\mathfrak{S})$ as its representative. The correspondent b^* of b in the basic ring A^*, constructed for $\mathfrak{S}$ as described in Theorem 36, is seen to coincide with the given set $\mathfrak{G}$. Hence the basic ring A^* contains every open set in $\mathfrak{S}$ as an element. If $\mathfrak{a}^*$ is the ideal of nowhere dense sets in A^* and if $\mathfrak{p}$ and $\mathfrak{q}$ are distinct points in $\mathfrak{C}'(\mathfrak{a}^*)$, there exists an element a^* in A^* such that $\mathfrak{C}(a^*)$ contains $\mathfrak{p}$ but not $\mathfrak{q}$. If b^* is the interior of a^* relative to $\mathfrak{S}$, then b^* is also in A^*. Furthermore, the relation $a^* \equiv b^*$ (mod $\mathfrak{a}^*$) implies the relation $\mathfrak{C}(a^*)\mathfrak{C}'(\mathfrak{a}^*)=\mathfrak{C}(b^*)\mathfrak{C}'(\mathfrak{a}^*)$. Hence $\mathfrak{C}(b^*)$ contains $\mathfrak{p}$ but not $\mathfrak{q}$. By the previous remarks, the open set b^* in $\mathfrak{S}$ corresponds to an element b in A through the relation $b^*=\mathfrak{G}(b)$ of Theorem 36. We can now conclude that the intersection of all sets $\mathfrak{C}(b^*)$ which contain $\mathfrak{p}\epsilon\mathfrak{C}'(\mathfrak{a}^*)$ and which correspond to sets b in A has no point in common with $\mathfrak{C}'(\mathfrak{a}^*)$ other than the given point. As we saw in the last part of the proof of Theorem 36, this intersection is a set $\mathfrak{F}^*$ belonging to $\mathfrak{f}^{**}$. To complete the proof of the present theorem we need only show that $\mathfrak{F}^*$ is a minimal set in $\mathfrak{f}^{**}$. We know that $\mathfrak{F}^*$ contains a minimal set $\mathfrak{F}_1^*$ in $\mathfrak{f}^{**}$: for, if $\mathfrak{F}$ is an $\mathfrak{X}(\mathfrak{S})$-set in the map $m(\mathfrak{S}, \mathfrak{B}(\mathfrak{S}), \mathfrak{X}(\mathfrak{S}))$, it contains a minimal set $\mathfrak{F}_1$; and its correspondent $\mathfrak{F}_1^*=f(\mathfrak{F}_1)$ is a minimal set belonging to $\mathfrak{f}^{**}$ and contained in $\mathfrak{F}^*$. Since the set $\mathfrak{F}^*$ is an $\mathfrak{X}^*$-set it must contain a point of $\mathfrak{C}'(\mathfrak{a}^*)$ by Theorem 34. Since $\mathfrak{F}^*$ has only the point $\mathfrak{p}$ in common with $\mathfrak{C}'(\mathfrak{a}^*)$, we see that $\mathfrak{F}_1^*$ also contains $\mathfrak{p}$. Now $\mathfrak{F}^*$ is by construction the set in $\mathfrak{f}^{**}$ mini-

mal with respect to the property of containing $\mathfrak{p}$. Hence $\mathfrak{F}^*$ coincides with $\mathfrak{F}_1^*$, and is minimal in $\mathfrak{F}^{**}$.

4. **Applications to the theory of extensions.** One of the interesting and difficult problems of general set-theoretic topology is the study of the extensions of a given space. The term "extension" is used here in the sense indicated by the following definition.

DEFINITION 13. *If a T_0-space $\mathfrak{Q}$ contains a subspace $\mathfrak{R}_{\mathfrak{Q}}$ equivalent to a given T_0-space $\mathfrak{R}$, then $\mathfrak{Q}$ is said to be an extension of $\mathfrak{R}$; and $\mathfrak{R}$ is said to be imbedded in $\mathfrak{Q}$ as the subspace $\mathfrak{R}_{\mathfrak{Q}}$. If $\mathfrak{R}_{\mathfrak{Q}}$ is a proper subset of $\mathfrak{Q}$, then $\mathfrak{Q}$ is said to be a proper extension of $\mathfrak{R}$; and if $\overline{\mathfrak{R}_{\mathfrak{Q}}}=\mathfrak{Q}$, the space $\mathfrak{Q}$ is said to be an immediate extention of $\mathfrak{R}$.*

The problem of extensions falls naturally into two distinct parts. If $\mathfrak{R}$ is imbedded in $\mathfrak{Q}$ as the subspace $\mathfrak{R}_{\mathfrak{Q}}$, then the subspace $\overline{\mathfrak{R}_{\mathfrak{Q}}}$ is evidently an immediate extension of $\mathfrak{R}$ and a closed subset of $\mathfrak{Q}$. Thus the determination of possible extensions of a given T_0-space $\mathfrak{R}$ involves, first, the determination of an immediate extension of $\mathfrak{R}$, and, secondly, the determination of a space in which this immediate extension can be imbedded as a closed subset. The second step is one which is obviously more arbitrary than the first, since the local structure of $\mathfrak{Q}$ at points "remote" from $\overline{\mathfrak{R}_{\mathfrak{Q}}}$ can in general be modified in a quite essential way without regard to the properties of $\mathfrak{R}$ or of $\mathfrak{R}_{\mathfrak{Q}}$. Thus the second step becomes most interesting when some additional requirement, say, of connectivity or dimensionality, is laid upon the space $\mathfrak{Q}$. The first step, on the other hand, appeals to the intuition as one which is intimately linked with the structure of the given space $\mathfrak{R}$. Here we shall confine our attention to the first step. As we proceed, we shall see in greater detail how the structure of $\mathfrak{R}$ determines that of its immediate extensions and how the theory of Boolean maps gives us a real insight into the problem under consideration.

In the course of our investigations we shall find it desirable to classify the immediate extensions of a given space. For convenience, we collect the appropriate definitions here.

DEFINITION 14. *An immediate extension $\mathfrak{Q}$ of a T_0-space $\mathfrak{R}$ is said to be a strict extension of $\mathfrak{R}$, if, when $\mathfrak{R}$ is imbedded in $\mathfrak{Q}$ as the subspace $\mathfrak{R}_{\mathfrak{Q}}$, the following property is verified: if $\mathfrak{G}$ is any open set in $\mathfrak{Q}$ and $\mathfrak{q}$ is any point in $\mathfrak{G}$, there exists an open set $\mathfrak{H}$ in $\mathfrak{Q}$ such that $\mathfrak{q}\in\mathfrak{H}\subset\mathfrak{G}$ and such that, whenever $\mathfrak{H}^*$ differs from $\mathfrak{H}$ by a nowhere dense set contained in $\mathfrak{R}'_{\mathfrak{Q}}$, the interior of $\mathfrak{H}^*$ is contained in $\mathfrak{G}$.*

DEFINITION 15. *An extension $\mathfrak{Q}$ of a T_0-space $\mathfrak{R}$ is said to be a T_1-extension*

of $\mathfrak{R}$ if, when $\mathfrak{R}$ is imbedded in $\mathfrak{Q}$ as the subspace $\mathfrak{R}_{\mathfrak{Q}}$, the relations $\mathfrak{q}\epsilon\mathfrak{Q}$, $\mathfrak{r}\epsilon\mathfrak{R}'_{\mathfrak{Q}}$, and $\mathfrak{q}\neq\mathfrak{r}$ imply $\mathfrak{r}\notin\{\mathfrak{q}\}^-$, $\mathfrak{q}\notin\{\mathfrak{r}\}^-$; in other words, that each of the points $\mathfrak{q}$, $\mathfrak{r}$ is contained in an open subset of $\mathfrak{Q}$ which does not contain the other.

DEFINITION 16. *An extension $\mathfrak{Q}$ of a T_0-space $\mathfrak{R}$ is said to be an H-extension of $\mathfrak{R}$ if, when $\mathfrak{R}$ is imbedded in $\mathfrak{Q}$ as the subspace $\mathfrak{R}_{\mathfrak{Q}}$, the relations $\mathfrak{q}\epsilon\mathfrak{Q}$, $\mathfrak{r}\epsilon\mathfrak{R}'_{\mathfrak{Q}}$, and $\mathfrak{q}\neq\mathfrak{r}$ imply the existence of two disjoint open subsets of $\mathfrak{Q}$ which contain $\mathfrak{q}$ and $\mathfrak{r}$ respectively.*

While the significance of Definitions 15 and 16 is evident, we may comment briefly on Definition 14. The property on which the latter definition is based means roughly that the points of $\mathfrak{R}'_{\mathfrak{Q}}$ are no more "densely" distributed in $\mathfrak{Q}$ than are those of $\mathfrak{R}_{\mathfrak{Q}}$. A strict extension of $\mathfrak{R}$ is therefore one in which the "new" points are not adjoined in too lavish a manner. The technical reasons for introducing the particular form of definition which has just been set forth, will be developed below.

An obvious method for constructing immediate extensions of a given T_0-space $\mathfrak{R}$ is based on the use of Boolean maps together with the concept of $\mathfrak{X}$-sets introduced in Definition 12. We obtain the following theorem.

THEOREM 38. *Let $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ be an arbitrary Boolean map of a T_0-space $\mathfrak{R}$ in a bicompact Boolean space $\mathfrak{B}$; and let $\mathfrak{Z}$ be any family of $\mathfrak{X}$-sets in $\mathfrak{B}$ containing the family $\mathfrak{X}$. Then under the topology of Theorems 14 and 23, $\mathfrak{Z}$ is a T_0-space which is an immediate extension of $\mathfrak{R}$. In order that $\mathfrak{Z}$ be a T_1-extension of $\mathfrak{R}$, it is necessary and sufficient that no set in $\mathfrak{Z}-\mathfrak{X}$ contain or be contained in any distinct set belonging to $\mathfrak{Z}$. In order that $\mathfrak{Z}$ be an H-extension of $\mathfrak{R}$, it is sufficient that no set in $\mathfrak{Z}-\mathfrak{X}$ have a point in common with any distinct set belonging to $\mathfrak{Z}$.*

If we consider $\mathfrak{Z}$ and $\mathfrak{X}$ as topological spaces, it is evident that $\mathfrak{X}$ is a subspace everywhere dense in $\mathfrak{Z}$: for, if $\mathfrak{z}_0$ is any element of $\mathfrak{Z}$ and $\mathfrak{G}$ is any open subset of $\mathfrak{B}$ containing $\mathfrak{z}_0$, the neighborhood of $\mathfrak{z}_0$ specified by the relation $\mathfrak{z}\subset\mathfrak{G}$ contains some element $\mathfrak{x}$ of the subspace $\mathfrak{X}$, by virtue of the fact that $\mathfrak{z}_0$ is an $\mathfrak{X}$-set. Since $\mathfrak{R}$ is equivalent to $\mathfrak{X}$, by the definition of a map, the space $\mathfrak{Z}$ is an immediate extension of $\mathfrak{R}$. The conditions for $\mathfrak{Z}$ to be a T_1- or an H-extension of $\mathfrak{R}$ are obtained automatically when one adjusts the argument used in Theorem 23 to the requirements of Definitions 15 and 16.

In order to show that in Theorem 38 we may restrict attention to the algebraic Boolean maps without any loss of generality, we must establish some algebraic preliminaries.

THEOREM 39. *If $\mathfrak{R}$ is a subspace of the T_0-space $\mathfrak{Q}$ such that $\mathfrak{R}^-=\mathfrak{Q}$ and*

if A is any subring of the complete basic ring $A_{\mathfrak{Q}}$, then the correspondence $a \rightarrow a\mathfrak{R}$ carries A homomorphically into a subring B of the complete basic ring $A_{\mathfrak{R}}$. If $\mathfrak{b}$ is the ideal in A defined by the relation $a\mathfrak{R}=0$, then $\mathfrak{b}$ consists of all nowhere dense sets which belong to A and are contained in $\mathfrak{R}'$; and B is an isomorph of $A/\mathfrak{b}$. The sets $a\mathfrak{R}$ in B which are nowhere dense relative to $\mathfrak{R}$ are precisely those for which a is nowhere dense relative to $\mathfrak{Q}$.

It is evident that the correspondence $a \rightarrow a\mathfrak{R}$ carries A homomorphically into a ring B of subsets of $\mathfrak{R}$; and that B is an isomorph of $A/\mathfrak{b}$ in accordance with R Theorem 43. If a does not belong to the ideal $\mathfrak{a}$ of nowhere dense sets in A, its interior $a'^{-'}$ is a non-void open subset of a. Since $\mathfrak{R}^{-}=\mathfrak{Q}$ implies $a'^{-'}\mathfrak{R} \neq 0$, the set $a\mathfrak{R}$ contains a non-void subset, $a'^{-'}\mathfrak{R}$, which is open relative to $\mathfrak{R}$. Thus $a\mathfrak{R}$ has interior points relative to $\mathfrak{R}$ and cannot be a nowhere dense subset of $\mathfrak{R}$. On the other hand, if a is nowhere dense relative to $\mathfrak{Q}$, we can show that $a\mathfrak{R}$ is nowhere dense relative to $\mathfrak{R}$. The closure of $a\mathfrak{R}$ relative to $\mathfrak{R}$ is $(a\mathfrak{R})^{-}\mathfrak{R}$, the complement of this set relative to $\mathfrak{R}$ is $(a\mathfrak{R})^{-'}\mathfrak{R}$, and the closure of the latter set relative to $\mathfrak{R}$ is $[(a\mathfrak{R})^{-'}\mathfrak{R}]^{-}\mathfrak{R}$. We therefore have to prove that $[(a\mathfrak{R})^{-'}\mathfrak{R}]^{-}\mathfrak{R}=\mathfrak{R}$. In view of the relations $(a\mathfrak{R})^{-}\mathfrak{R} \subset a^{-}\mathfrak{R}$, $(a\mathfrak{R})^{-'}\mathfrak{R} \supset a^{-'}\mathfrak{R}$, it is evidently sufficient to prove that $(a^{-'}\mathfrak{R})^{-}=\mathfrak{Q}$. Now our assumption that a is nowhere dense means that $a^{-'-}=\mathfrak{Q}$. If we denote by $\mathfrak{G}$ the open set $a^{-'}$ in $\mathfrak{Q}$, we can complete our demonstration by deducing $(\mathfrak{G}\mathfrak{R})^{-}=\mathfrak{Q}$ from the known relations $\mathfrak{G}^{-}=\mathfrak{R}^{-}=\mathfrak{Q}$, $\mathfrak{G}'^{-}=\mathfrak{G}'$. Since $\mathfrak{G}\mathfrak{R} \subset (\mathfrak{G}\mathfrak{R})^{-}$, we have $(\mathfrak{G}\mathfrak{R})(\mathfrak{G}\mathfrak{R})^{-'}=0$, $\mathfrak{R} \subset [\mathfrak{G}(\mathfrak{G}\mathfrak{R})^{-'}]'=\mathfrak{G}' \cup (\mathfrak{G}\mathfrak{R})^{-}$, $\mathfrak{Q}=\mathfrak{R}^{-} \subset [\mathfrak{G}' \cup (\mathfrak{G}\mathfrak{R})^{-}]^{-}=\mathfrak{G}' \cup (\mathfrak{G}\mathfrak{R})^{-}$, $\mathfrak{G} \subset (\mathfrak{G}\mathfrak{R})^{-}$, $\mathfrak{Q}=\mathfrak{G}^{-} \subset (\mathfrak{G}\mathfrak{R})^{--}=(\mathfrak{G}\mathfrak{R})^{-}$, and hence $(\mathfrak{G}\mathfrak{R})^{-}=\mathfrak{Q}$, as we desired to show. Thus we see that, when $a \epsilon A$, the set $a\mathfrak{R}$ is nowhere dense relative to $\mathfrak{R}$ if and only if $a \epsilon \mathfrak{a}$. It follows that the relation $a \equiv a'^{-'}$ (mod $\mathfrak{a}$) in A implies the relation $a\mathfrak{R} \equiv a'^{-'}\mathfrak{R}$ (mod $\mathfrak{a}_{\mathfrak{R}}$), where $\mathfrak{a}_{\mathfrak{R}}$ is the ideal of all nowhere dense sets in $A_{\mathfrak{R}}$. Since $a'^{-'}\mathfrak{R}$ is open relative to $\mathfrak{R}$, it is clear that, by definition, $a\mathfrak{R}$ belongs to $A_{\mathfrak{R}}$. The various further properties of B and $\mathfrak{b}$ are now evident.

THEOREM 40. *Let $\mathfrak{R}$, $\mathfrak{Q}$, A, B, and $\mathfrak{b}$ have the same meanings as in the preceding theorem; let B have the property that it is a basic ring for $\mathfrak{R}$; let $m(\mathfrak{Q}, \mathfrak{C}(A), \mathrm{Z})$ and $m(\mathfrak{R}, \mathfrak{C}(B), \Upsilon)$ be the algebraic maps determined by A and B respectively; and let $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ be the map obtained by suppressing from the family Z in $m(\mathfrak{Q}, \mathfrak{C}(A), \mathrm{Z})$ those members which correspond to points of $\mathfrak{R}'$. Then a set in $\mathfrak{C}(A)$ is a Z-set if and only if it is an $\mathfrak{X}$-set. If A contains a basis for $\mathfrak{Q}$, then B contains a basis for $\mathfrak{R}$ and is a basic ring for $\mathfrak{R}$; and the map $m(\mathfrak{R}, \mathfrak{C}(B), \Upsilon)$ is equivalent to the map obtained from $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ by the suppression of the open set $\mathfrak{C}(\mathfrak{b})$. Under the same condition on A, the set $\mathfrak{C}(\mathfrak{b})$*

is a set of redundancy in the map $m(\mathfrak{Q}, \mathfrak{E}(A), \mathrm{Z})$ *if and only if* $\mathfrak{R}$, $\mathfrak{Q}$, *and* $\mathfrak{b}$ *are related in the following manner*:

(P) *if* $\mathfrak{G}$ *is any open set in* $\mathfrak{Q}$ *and* $\mathfrak{q}$ *is any point in* $\mathfrak{G}$, *there exists an open set* $\mathfrak{H}$ *in* $\mathfrak{Q}$ *such that* $\mathfrak{q}\epsilon\mathfrak{H} \subset \mathfrak{G}$ *and such that, whenever* $\mathfrak{H}^* \equiv \mathfrak{H}$ (mod $\mathfrak{b}$), *the interior of* $\mathfrak{H}^*$ *is contained in* $\mathfrak{G}$.

The property (P) *holds whenever* $\mathfrak{Q}$ *is a strict extension of* $\mathfrak{R}$; *and* $\mathfrak{Q}$ *is a strict extension of* $\mathfrak{R}$ *whenever the property* (P) *holds for the case where A is the complete basic ring for* $\mathfrak{Q}$. *When* $\mathfrak{E}(\mathfrak{b})$ *is a set of redundancy in accordance with the foregoing conditions, its suppression from the map* $m(\mathfrak{Q}, \mathfrak{E}(A), \mathrm{Z})$ *yields a map equivalent to* $m(\mathfrak{Q}, \mathfrak{E}(B), \mathcal{W})$, *where* $\mathcal{W}$ *is a family of* Υ*-sets containing* Υ *as a subfamily in the map* $m(\mathfrak{R}, \mathfrak{E}(B), \Upsilon)$.

If $\mathfrak{F}$ is any Z-set in $\mathfrak{E}(A)$ and $\mathfrak{E}(a) \supset \mathfrak{F}$, then there exists a set $\mathfrak{Z}$ in Z such that $\mathfrak{E}(a) \supset \mathfrak{Z}$. If $\mathfrak{q}$ is the correspondent of this set in $\mathfrak{Q}$, then $\mathfrak{q}\epsilon a'^{-\prime}$ in accordance with Theorem 28. Since $a'^{-\prime}$ is open in $\mathfrak{Q}$, the fact that $\mathfrak{R}^- = \mathfrak{Q}$ implies the existence of a point $\mathfrak{r}$ in $\mathfrak{R}$ which belongs to $a'^{-\prime}$. If $\mathfrak{X}$ is the set in $X \subset \mathrm{Z}$ corresponding to the point $\mathfrak{r}$, then $\mathfrak{E}(a) \supset \mathfrak{X}$. Hence the set $\mathfrak{F}$ is an X-set in accordance with Definition 12. Since $X \subset \mathrm{Z}$, Theorem 30 shows that every X-set is a Z-set. Thus the X-sets are identical with the Z-sets.

If A contains a basis in $\mathfrak{Q}$, then B contains a basis in $\mathfrak{R}$: for the sets in any basis in $\mathfrak{Q}$ intersect $\mathfrak{R}$ in a basis for $\mathfrak{R}$. Thus, in this case, B is a basic ring for $\mathfrak{R}$. In view of the isomorphism between B and the quotient-ring $A/\mathfrak{b}$, the closed set $\mathfrak{E}'(\mathfrak{b})$ in $\mathfrak{E}(A)$ is a Boolean space representing B in accordance with Theorem 4; and $\mathfrak{E}'(\mathfrak{b})$ and $\mathfrak{E}(B)$ are topologically equivalent in such a manner that the relation $b = a\mathfrak{R}$ implies the correspondence of the sets $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$ and $\mathfrak{E}(b)$ under the equivalence. In order that $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(b)$, where $\mathfrak{r}\epsilon\mathfrak{R}$ and $b\epsilon B$, it is necessary and sufficient that $\mathfrak{r}$ be an interior point of b relative to $\mathfrak{R}$. In order that $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$, where $\mathfrak{r}\epsilon\mathfrak{R}$ and $a\epsilon A$, it is necessary and sufficient that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a)$ or, equivalently, that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(c)$, where c is an element of A such that $c \equiv a$ (mod $\mathfrak{b}$). A proof of this assertion runs as follows: the relations $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a)$ and $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a) \subset \mathfrak{E}(\mathfrak{b})$ are equivalent; since $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a)$ is a closed set in the bicompact space $\mathfrak{E}(A)$, a familiar argument shows that the second of these relations can hold if and only if there exists an element d in $\mathfrak{b}$ such that $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a) \subset \mathfrak{E}(d)$; if we take $c = a \vee d \equiv a$ (mod $\mathfrak{b}$), we have $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(a) \cup \mathfrak{E}(d) = \mathfrak{E}(c)$; and, on the other hand, if $c \equiv a$ (mod $\mathfrak{b}$) and $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(c)$, we have $d = a'c\epsilon\mathfrak{b}$, $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(a) \subset \mathfrak{E}(c)\mathfrak{E}'(a) = \mathfrak{E}(a'c) = \mathfrak{E}(d)$. With the help of the indicated characterizations of the relations $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(b)$, $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$, we shall now show that the equivalence between $\mathfrak{E}'(\mathfrak{b})$ and $\mathfrak{E}(B)$ carries $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r})$ into $\mathfrak{Y}(\mathfrak{r})$. First, if we have $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$, we find an element c of the kind described

above, noting that $c \equiv a$ (mod $\mathfrak{b}$) implies $b = a\mathfrak{R} = c\mathfrak{R}$ by the definition of the ideal $\mathfrak{b}$, and that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(c)$ implies $\mathfrak{r} \epsilon c'^{-\prime}$. Since $\mathfrak{r}$ is in the set $c'^{-\prime}\mathfrak{R}$ which is contained in the interior of $b = c\mathfrak{R}$ relative to $\mathfrak{R}$, we see that $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(b)$. Now $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r})$ is the intersection of all the sets $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$ containing it; its image is therefore the intersection of all the corresponding sets $\mathfrak{E}(b)$ in $\mathfrak{E}(B)$, where $b = a\mathfrak{R}$; and hence its image contains $\mathfrak{Y}(\mathfrak{r})$. Secondly, if we have $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{E}(b)$ or, equivalently, if $\mathfrak{r}$ is in the interior of b relative to $\mathfrak{R}$, our assumption that A contains a basis for $\mathfrak{Q}$ leads to the result that there exists a set $d = c\mathfrak{R}$, where c is an open set belonging to A and d is therefore open relative to $\mathfrak{R}$, with the properties $\mathfrak{r} \epsilon d < b$. Hence the relations $\mathfrak{r} \epsilon c = c'^{-\prime}$, $\mathfrak{E}(d) \subset \mathfrak{E}(b)$ imply the relations $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(c)$, $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(c) \subset \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$, where $d = c\mathfrak{R}$ and $b = a\mathfrak{R}$. We conclude that $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$. Now since $\mathfrak{Y}(\mathfrak{r})$ is the intersection of all the sets $\mathfrak{E}(b)$ containing it, where $b \epsilon B$, its image in $\mathfrak{E}'(\mathfrak{b})$ is the intersection of all the corresponding sets $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(a)$, where $b = a\mathfrak{R}$; and its image therefore contains the set $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r})$. Combining these results, we see that the sets $\mathfrak{E}'(\mathfrak{b})\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{Y}(\mathfrak{r})$ are images of one another, as we wished to prove. Accordingly, the suppression of the open set $\mathfrak{E}(\mathfrak{b})$ from the map $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X})$ yields a map equivalent to $m(\mathfrak{R}, \mathfrak{E}(B), \mathfrak{Y})$. It follows that $\mathfrak{E}(\mathfrak{b})$ is a set of redundancy in harmony with Definitions 8 and 9.

It is now possible to analyze the removal of the set $\mathfrak{E}(\mathfrak{b})$ from the map $m(\mathfrak{Q}, \mathfrak{E}(A), \mathfrak{Z})$. Applying the criterion given in Theorem 18, we see that $\mathfrak{E}(\mathfrak{b})$ is a set of redundancy if and only if, whenever $\mathfrak{E}(a) \supset \mathfrak{Z}_0$, there exists a set a_0 in A such that $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}_0$ while $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$ implies $\mathfrak{E}(a) \supset \mathfrak{Z}$. Let us consider the sets in A which are congruent (mod $\mathfrak{b}$) to such a set a_0. If c is any such set, we have $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) = \mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(c)$ since $\mathfrak{E}(a_0)\Delta\mathfrak{E}(c) = \mathfrak{E}(a_0 + c) \subset \mathfrak{E}(\mathfrak{b})$. As we proved in the preceding paragraph, the relation $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$ is equivalent to the relation $\mathfrak{E}(c) \supset \mathfrak{Z}$ for some such set c. Hence it is possible to choose a_0 so that $\mathfrak{E}(a_0) \supset \mathfrak{Z}_0$ while $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$ implies $\mathfrak{E}(a) \supset \mathfrak{Z}$: for if a_0 does not have the first property we can replace it by a congruent set which does, and this substitution does not affect the significance of the inclusion $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$. The point $\mathfrak{q}$ corresponding to $\mathfrak{Z}_0$ is then interior to a_0 as well as to a. In order that the inclusion $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$ imply $\mathfrak{E}(a) \supset \mathfrak{Z}$, it is necessary and sufficient that $c \equiv a_0$ (mod $\mathfrak{b}$) imply $c'^{-\prime} < a'^{-\prime}$. We begin with the necessity of this condition. The sets $c'^{-\prime}$, $a'^{-\prime}$ are the subsets of $\mathfrak{Q}$ specified by the relations $\mathfrak{E}(c) \supset \mathfrak{Z}$, $\mathfrak{E}(a) \supset \mathfrak{Z}$. Hence, from the assumptions $c \equiv a_0$ (mod $\mathfrak{b}$), $\mathfrak{E}(c) \supset \mathfrak{Z}$, we can deduce the relations $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) = \mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(c) \supset \mathfrak{Z}$, $\mathfrak{E}(a) \supset \mathfrak{Z}$, and $c'^{-\prime} < a'^{-\prime}$. We pass then to the sufficiency. If $\mathfrak{E}(\mathfrak{b}) \cup \mathfrak{E}(a_0) \supset \mathfrak{Z}$, there exists a set c in A congruent to a_0 (mod $\mathfrak{b}$) such that $\mathfrak{E}(c) \supset \mathfrak{Z}$; and the assumption that $c \equiv a_0$ (mod $\mathfrak{b}$) implies $c'^{-\prime} < a'^{-\prime}$ then leads to the relation $\mathfrak{E}(a) \supset \mathfrak{Z}$. In view of the preceding discussion, we see that $\mathfrak{E}(\mathfrak{b})$ is a set of redundancy

if and only if, whenever a is a set in A and $\mathfrak{q}$ an interior point of a, there exists a set a_0 in A which contains $\mathfrak{q}$ as an interior point and which has the property that, for c in A, the relation $c \equiv a_0$ (mod $\mathfrak{b}$) implies $c'^{-'} < a'^{-'}$. It is evident that if A contains a basis this condition can be replaced by the requirement that $\mathfrak{Q}$, $\mathfrak{R}$, and $\mathfrak{b}$ have the property (P) stated in the theorem. First, let us deduce (P) from the condition just given. If $\mathfrak{G}$ is any open set in $\mathfrak{Q}$ and $\mathfrak{q}$ any point of $\mathfrak{G}$, we choose a in A so that $\mathfrak{q} \epsilon a'^{-'} \subset \mathfrak{G}$; we may, of course, take a as an open set if we wish. We then determine a_0 and choose $\mathfrak{H}$ as an open set in A so that $\mathfrak{q} \epsilon \mathfrak{H} \subset a_0'^{-'}$. Then, if $\mathfrak{H}^* \equiv \mathfrak{H}$ (mod $\mathfrak{b}$), we see that $\mathfrak{H}^*$ belongs to A and is contained in the set $c = a_0 \cup \mathfrak{H}'\mathfrak{H}^* \equiv a_0$ (mod $\mathfrak{b}$), where c also is in A. Hence we have $(\mathfrak{H}^*)'^{-'} \subset c'^{-'} \subset a'^{-'} \subset \mathfrak{G}$. Thus (P) is verified. On the other hand, if (P) holds, we apply it, taking $\mathfrak{G} = a'^{-'}$ with a in A and $\mathfrak{q} \epsilon a'^{-'}$, so as to determine an open set $\mathfrak{H}$ containing $\mathfrak{q}$ and possessing the other indicated properties. We then choose a_0 as an open set in A so that $\mathfrak{q} \epsilon a_0 \subset \mathfrak{H}$. If $c \equiv a_0$ (mod $\mathfrak{b}$), we see that c belongs to A and is contained in $\mathfrak{H}^* = \mathfrak{H} \cup a_0' c \equiv \mathfrak{H}$ (mod $\mathfrak{b}$). Hence we have $c'^{-'} < (\mathfrak{H}^*)'^{-'} \subset \mathfrak{G} = a'^{-'}$, as we wished to prove.

On comparison with Definition 14, it is evident that the property (P) holds whenever $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. On the other hand, if A is the complete basic ring of $\mathfrak{Q}$, the ideal $\mathfrak{b}$ consists of all the nowhere dense subsets of $\mathfrak{R}'$; and, if the property (P) holds in this case, $\mathfrak{Q}$ must therefore be a strict extension of $\mathfrak{R}$. The remaining statements of the theorem are obvious consequences of the results already obtained.

On combining Theorems 39 and 40, we see that the construction of extensions of a given space $\mathfrak{R}$ can be carried out in the following way:

THEOREM 41. *If $\mathfrak{R}$ is an arbitrary T_0-space and $A_{\mathfrak{R}}$ is the complete basic ring of $\mathfrak{R}$, then every immediate extension of $\mathfrak{R}$ can be found by the following construction: the algebraic map $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{R}}), \Upsilon)$ is constructed, the space $\mathfrak{E}(A_{\mathfrak{R}})$ is then imbedded in any suitable bicompact Boolean space $\mathfrak{B}$, and each set $\mathfrak{Y}$ is enlarged by the adjunction of points in $\mathfrak{B} - \mathfrak{E}(A_{\mathfrak{R}})$ to form a set $\mathfrak{X}$ in such a way that the resulting family X in $\mathfrak{B}$ provides a map $m(\mathfrak{R}, \mathfrak{B}, X)$; and then the construction of Theorem 39 is applied to the latter map. In this procedure the following specializations are possible:*

(1) *every strict extension of $\mathfrak{R}$ can be obtained under the conditions $\mathfrak{B} = \mathfrak{E}(A_{\mathfrak{R}}), X = \Upsilon$;*
(2) *there is no loss of generality in supposing that the enlarged sets $\mathfrak{X}$ corresponding to sets $\mathfrak{Y}$ are disjoint whenever the latter are disjoint;*
(3) *every immediate H-extension can be obtained under the restriction described in (2) and the further restriction that no set $\mathfrak{Z}$ in the family*

$\mathfrak{Z}-\mathfrak{X}$ *have a point in common with any distinct set in the family* $\mathfrak{Z}$, *the notations being those of Theorem* 39;

(4) *every strict H-extension can be obtained under the conditions described in* (1) *and* (3), *that in* (2) *being trivial.*

The foregoing construction can be applied to any algebraic map $m(\mathfrak{R}, \mathfrak{E}(A), \Upsilon)$, *but does not necessarily provide all extensions of* $\mathfrak{R}$ *when* $A \neq A_{\mathfrak{R}}$.

If $\mathfrak{Q}$ is any immediate extension of $\mathfrak{R}$, we can apply Theorem 40 using the complete basic ring $A_{\mathfrak{Q}}$ of $\mathfrak{Q}$. The corresponding basic ring B then coincides with the complete basic ring $A_{\mathfrak{R}}$ for $\mathfrak{R}$, as we see by reference to Theorem 39. Indeed, it is evident that B contains every subset of $\mathfrak{R}$ which is open relative to $\mathfrak{R}$, since $A_{\mathfrak{Q}}$ contains every open subset of $\mathfrak{Q}$. Furthermore, if $\mathfrak{F}$ is a subset of $\mathfrak{R}$ nowhere dense relative to $\mathfrak{R}$, we see that $\mathfrak{F}$ is in $A_{\mathfrak{Q}}$ and hence also in $A_{\mathfrak{R}}$: for the assumed relation $(\mathfrak{F}^{-\prime}\mathfrak{R})^{-}\mathfrak{R}=\mathfrak{R}$ implies $\mathfrak{Q}=\mathfrak{R}^{-}\subset(\mathfrak{F}^{-\prime}\mathfrak{R})^{-}\subset\mathfrak{F}^{-\prime-}\mathfrak{R}^{-}=\mathfrak{F}^{-\prime-}$, so that $\mathfrak{F}$ is nowhere dense relative to $\mathfrak{Q}$. According to Theorem 24, it follows that B contains $A_{\mathfrak{R}}$; but, as a subring of $A_{\mathfrak{R}}$, B coincides with $A_{\mathfrak{R}}$. Hence the map $m(\mathfrak{Q}, \mathfrak{E}(A_{\mathfrak{Q}}), \mathfrak{Z})$ is related to the map $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{R}}), \Upsilon)$ in the manner indicated in Theorem 40 and described from another point of view in the constructive program stated above. Thus our construction is capable of providing all possible immediate extensions of $\mathfrak{R}$. If $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$, we know from Theorem 40 that the set $\mathfrak{E}(\mathfrak{b})=\mathfrak{E}(A_{\mathfrak{Q}})-\mathfrak{E}(A_{\mathfrak{R}})$ can be suppressed from the map $m(\mathfrak{Q}, \mathfrak{E}(A_{\mathfrak{Q}}), \mathfrak{Z})$ so as to yield a map $m(\mathfrak{Q}, \mathfrak{E}(A_{\mathfrak{R}}), \mathfrak{W})$ with $\mathfrak{W}\supset\Upsilon$. Thus we obtain all strict extensions of $\mathfrak{R}$ under the special conditions stated in (1). If we now refer to the proof of Theorem 28, we see that two distinct points in $\mathfrak{R}$ (or in $\mathfrak{Q}$) have the H-separation property, namely, the property of belonging respectively to two disjoint open sets, if and only if their representative sets $\mathfrak{Y}$ (or $\mathfrak{Z}$) are disjoint. We observe that two distinct points in $\mathfrak{R}$ have the H-separation property relative to $\mathfrak{R}$ if and only if they have that property also relative to $\mathfrak{Q}$: for the open sets in $\mathfrak{R}$ are precisely the sets $\mathfrak{G}\mathfrak{R}$ where $\mathfrak{G}$ is open in $\mathfrak{Q}$; and the relations $(\mathfrak{G}_1\mathfrak{G}_2)\mathfrak{R}=(\mathfrak{G}_1\mathfrak{R})(\mathfrak{G}_2\mathfrak{R})=0$ and $\mathfrak{G}_1\mathfrak{G}_2=0$ are equivalent by virtue of the fact that $\mathfrak{R}^{-}=\mathfrak{Q}$. Hence the sets $\mathfrak{X}$ in $m(\mathfrak{Q}, \mathfrak{E}(A_{\mathfrak{Q}}), \mathfrak{Z})$ or $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{Q}}), \mathfrak{X})$ corresponding to two points in $\mathfrak{R}$ are disjoint if and only if these points have the H-separation property relative to $\mathfrak{R}$; that is, if and only if the corresponding sets $\mathfrak{Y}$ in $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{R}}), \Upsilon)$ are disjoint. Thus we conclude that (2) and (3) are valid, the latter in accordance with Definition 16. Reviewing the preceding discussion, we see also that (4) is valid.

There are two general comments to be made on the construction described above. First, the imbedding of a Boolean space in another offers no difficulty: for, by reference to Chapter I, §3, we see that this can be accom-

plished, in all possible ways, by imbedding the given Boolean space as a set $\mathfrak{F}$ in one of the universal Boolean spaces $\mathfrak{B}_{\mathfrak{c}}$ of sufficiently great character $\mathfrak{c}$, and then retaining some Boolean subspace of $\mathfrak{B}_{\mathfrak{c}}$ which contains $\mathfrak{F}$. The whole process can, of course, be expressed in purely algebraic terms. Secondly, the construction of the family $\mathfrak{X}$ from the family $\mathfrak{Y}$ has not been analyzed in a precise way. We see in fact that there is a significant distinction between the concepts of immediate and of strict extensions. The strict extensions of a T_0-space $\mathfrak{R}$ are obviously quite closely bound by the topological structure of $\mathfrak{R}$, while the immediate extensions are related to $\mathfrak{R}$ in a somewhat vague manner.

The final statement of the theorem does not require elaboration.

As an immediate corollary of Theorem 41, we have the following result:

THEOREM 42. *Every immediate T_1-extension of a T_1-space is a T_1-space; and every immediate H-extension of an H-space is an H-space.*

We discuss only the second statement of the theorem, leaving the quite similar proof of the first part to the reader. If the T_0-space $\mathfrak{Q}$ is an H-extension of the H-space $\mathfrak{R}$, then we see, as in the proof of Theorem 41, that no set in the family $\mathfrak{Z}-\mathfrak{X}$ has a point in common with any distinct set in $\mathfrak{Z}$; and that the sets in $\mathfrak{X}$, like those in $\mathfrak{Y}$, are mutually disjoint. Hence the sets belonging to $\mathfrak{Z}$ are all disjoint; and $\mathfrak{Q}$ is therefore an H-space.

We shall now proceed to consider, with the help of the preceding general theory, several more specific problems concerning extensions. These problems all cluster about the following general definition:

DEFINITION 17. *A T_0-space $\mathfrak{R}$ is said to be absolutely closed with respect to a particular type of extension if it has no proper extension of that type.*

The first results which we report are well known and almost trivial.

THEOREM 43. *The only T_0-space which is absolutely closed with respect to immediate extension is the void space.*

Using the construction of Theorem 41, we map a non-void T_0-space $\mathfrak{R}$ in the space $\mathfrak{E}(A_{\mathfrak{R}})$, imbed $\mathfrak{E}(A_{\mathfrak{R}})$ as a proper subset in a bicompact Boolean space $\mathfrak{B}$, take $\mathfrak{X}=\mathfrak{Y}$, and obtain $\mathfrak{Z}$ by adjoining to $\mathfrak{X}$ a single $\mathfrak{X}$-set which has at least one point in common with $\mathfrak{B}-\mathfrak{E}(A_{\mathfrak{R}})$. The determination of such an $\mathfrak{X}$-set is easily carried out with the help of Theorem 34. Under the usual topology, $\mathfrak{Z}$ is an immediate T_0-extension of $\mathfrak{R}$. In any T_0-space the void set is closed, $0^-=0$. Hence the void T_0-space can be imbedded in any T_0-space and is closed therein. We see therefore that it has no proper immediate T_0-extension.

We can, however, obtain a stronger result than this.

THEOREM 44. *The only T_0-space which is absolutely closed with respect to strict T_0-extension is the void space. Indeed, every non-void T_0-space $\mathfrak{R}$ becomes, by suitable adjunction of a single point, a bicompact T_0-space which is a strict T_0-extension of $\mathfrak{R}$.*

Let $\mathfrak{Q}$ be the class obtained by adjoining a single point ξ to the space $\mathfrak{R}$. In $\mathfrak{Q}$, let $\mathfrak{F}$ be the class comprising the following subsets of $\mathfrak{Q}$: (1) the void set; (2) the sets $\mathfrak{F} \cup \{\xi\}$, where $\mathfrak{F}$ is closed in $\mathfrak{R}$. Then the class $\mathfrak{F}$ has the following properties: (1) the finite union and the arbitrary intersection of sets in $\mathfrak{F}$ are in $\mathfrak{F}$; (2) the void set and the set $\mathfrak{Q}$ are both in $\mathfrak{F}$. Hence we can introduce in $\mathfrak{Q}$ a closure operation such that the closed sets are precisely the sets belonging to $\mathfrak{F}$.* Since this closure operation has the properties that $0^- = 0$ and $\{\mathfrak{p}\}^- = \{\mathfrak{q}\}^-$ implies $\mathfrak{p} = \mathfrak{q}$, the space $\mathfrak{Q}$ is a T_0-space. In fact, $0^- = 0$ is trivial, and the second property is established as follows: if $\mathfrak{p} \neq \mathfrak{q}$ and $\mathfrak{p} \in \mathfrak{R}$, $\mathfrak{q} \in \mathfrak{R}$, then $\{\mathfrak{p}\}^-$ and $\{\mathfrak{q}\}^-$ are obtained by adjoining ξ to the corresponding closures, relative to $\mathfrak{R}$, of $\{\mathfrak{p}\}$ and $\{\mathfrak{q}\}$, so that $\{\mathfrak{p}\}^- \neq \{\mathfrak{q}\}^-$; and, if $\mathfrak{p} \in \mathfrak{R}$, then $\{\mathfrak{p}\}^-$ contains $\{\xi\}^- = \{\xi\}$ as a proper subset, the set $\{\xi\} = 0 \cup \{\xi\}$ being closed by definition. Moreover, the space $\mathfrak{Q}$ is bicompact. Indeed, if an open set $\mathfrak{G}$ in $\mathfrak{Q}$ contains ξ, then $\mathfrak{G}'$ is closed and does not contain ξ, so that $\mathfrak{G}' = 0$ and $\mathfrak{G} = \mathfrak{Q}$. Hence any family of open sets which covers $\mathfrak{Q}$ must contain a subfamily consisting of $\mathfrak{Q}$ alone, which already covers $\mathfrak{Q}$.

It remains for us to prove that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. The closed sets in $\mathfrak{R}$ are precisely the sets $\mathfrak{F}\mathfrak{R}$ where $\mathfrak{F}$ is closed in $\mathfrak{Q}$, so that $\mathfrak{Q}$ contains $\mathfrak{R}$ as a relative subspace. The closure of $\mathfrak{R}$ in $\mathfrak{Q}$ is the intersection of all closed subsets of $\mathfrak{Q}$ which contain $\mathfrak{R}$; but, since the only such set is $\mathfrak{Q}$ itself, we have $\mathfrak{R}^- = \mathfrak{Q}$. Hence $\mathfrak{Q}$ is an immediate extension of $\mathfrak{R}$. The only subsets of $\mathfrak{R}'$ are 0 and $\{\xi\}$. Since $0^{-\prime-} = 0'^- = \mathfrak{Q}^- = \mathfrak{Q}$, $\{\xi\}^{-\prime-} = \{\xi\}'^- = \mathfrak{R}^- = \mathfrak{Q}$, both sets are nowhere dense in $\mathfrak{Q}$. Thus the only way of modifying an open subset $\mathfrak{G}$ of $\mathfrak{Q}$ by operating with the nowhere dense subsets of $\mathfrak{R}'$ is to suppress or adjoin ξ. Hence we have to consider two cases under Definition 16: first, the case where $\mathfrak{G}$ contains ξ and this point is suppressed; and, second, the case where $\mathfrak{G}$ does not contain ξ and this point is adjoined. If $\mathfrak{G}$ contains ξ, then $\mathfrak{G} = \mathfrak{Q}$ and the open set $\mathfrak{G} - \{\xi\}$ is contained in $\mathfrak{G}$. If $\mathfrak{G}$ does not contain ξ, the interior of $\mathfrak{G} \cup \{\xi\}$ coincides with $\mathfrak{G}$: for $(\mathfrak{G} \cup \{\xi\})'^{-\prime} = (\mathfrak{G}'\mathfrak{R})^{-\prime} = (\mathfrak{G}'\mathfrak{R} \cup \{\xi\})' = (\mathfrak{G} \cup \mathfrak{R}')\mathfrak{R} = \mathfrak{G}$ since $\mathfrak{G}'\mathfrak{R}$ is closed in $\mathfrak{R}$ and obviously has $\mathfrak{G}'\mathfrak{R} \cup \{\xi\}$, a closed set in $\mathfrak{Q}$, as its closure relative to $\mathfrak{Q}$. Thus we see that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$ in harmony with Definition 16.

We have also the following fact.

* AH, p. 41, Satz VI.

THEOREM 45. *The only T_0-spaces which are absolutely closed with respect to immediate T_1-extension are the finite T_0-spaces.*

While it would no doubt be instructive to discuss this assertion by means of the general mapping theory, the argument is quite involved. It is therefore simpler to appeal to results given by Alexandroff and Hopf.† Their construction shows how to adjoin a single point to an infinite T_0-space so as to obtain an extension. It is easily verified that this extension is actually an immediate T_1-extension. On the other hand, if $\mathfrak{Q}$ is a T_1-extension of a finite T_0-space $\mathfrak{R}$, we can show that $\mathfrak{R}^- = \mathfrak{R}$ in $\mathfrak{Q}$ and can then conclude that $\mathfrak{R}$ has no proper immediate T_1-extension. In fact if $\mathfrak{p} \epsilon \mathfrak{R}$ and $\mathfrak{q} \epsilon \mathfrak{R}'$, we see that there exists an open set $\mathfrak{G}$ such that $\mathfrak{p} \epsilon \mathfrak{G}'$, $\mathfrak{q} \epsilon \mathfrak{G}$. Hence $\{\mathfrak{p}\}^- \subset \mathfrak{G}'^- = \mathfrak{G}'$ in $\mathfrak{Q}$; and we must have $\{\mathfrak{p}\}^- \subset \mathfrak{R}$, since $\mathfrak{q}$ is arbitrary in $\mathfrak{R}'$. If the points of $\mathfrak{R}$ are $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$, we therefore have $\mathfrak{R}^- = \{\mathfrak{p}_1\}^- \cup \cdots \cup \{\mathfrak{p}_n\}^- \subset \mathfrak{R}$, $\mathfrak{R}^- = \mathfrak{R}$, as we wished to prove.

In contrast with the foregoing results we shall now establish the existence of less trivial T_0-spaces which are absolutely closed with respect to strict T_1-extension.

THEOREM 46. *In an arbitrary infinite class $\mathfrak{R}$, let a closure operation be defined as follows*: (1) *if $\mathfrak{F}$ is a finite subset of $\mathfrak{R}$, then $\mathfrak{F}^- = \mathfrak{F}$*; (2) *if $\mathfrak{F}$ is an infinite subset of $\mathfrak{R}$, then $\mathfrak{F}^- = \mathfrak{R}$. Then $\mathfrak{R}$ is a bicompact T_1-space which is absolutely closed with respect to strict T_1-extension.*

It is easily verified that the indicated closure operation has the properties $\mathfrak{F}^- \supset \mathfrak{F}$, $\mathfrak{F}^{--} = \mathfrak{F}^-$, $(\mathfrak{F}_1 \cup \mathfrak{F}_2)^- = \mathfrak{F}_1^- \cup \mathfrak{F}_2^-$, $0^- = 0$, $\{\mathfrak{r}\}^- = \{\mathfrak{r}\}$, so that $\mathfrak{R}$ is a T_1-space. Moreover, $\mathfrak{R}$ is not an H-space. Indeed the only closed sets in $\mathfrak{R}$ are $\mathfrak{R}$ and its finite subsets, the only open sets in $\mathfrak{R}$ are the void set and the subsets differing from $\mathfrak{R}$ by finite sets, and any two non-void open sets must therefore have points in common. The nowhere dense sets in $\mathfrak{R}$ are precisely the finite subsets of $\mathfrak{R}$: for $\mathfrak{F}^{-\prime -} = \mathfrak{R}$ if and only if $\mathfrak{F}^{-\prime}$ is infinite, $\mathfrak{F}^-$ finite, and $\mathfrak{F}$ finite and equal to $\mathfrak{F}^-$. The complete basic ring $A_{\mathfrak{R}}$ for $\mathfrak{R}$ is thus seen to consist of the finite subsets of $\mathfrak{R}$ and their complements. In order to determine the Boolean space $\mathfrak{E}(A_{\mathfrak{R}})$, we first introduce a new topology in the class $\mathfrak{R}$, obtaining a space $\mathfrak{R}^*$. The closure operation is defined for this purpose by setting $\mathfrak{F}^- = \mathfrak{F}$ for every subset $\mathfrak{F}$ of $\mathfrak{R}$. In $\mathfrak{R}^*$ every set is both closed and open; in particular, the one-element sets constitute a basis in $\mathfrak{R}^*$. It is therefore evident that $\mathfrak{R}^*$ is a non-bicompact Boolean space in which the bicompact subspaces are precisely the finite subsets. Thus $\mathfrak{R}^*$ is a representative of the Boolean ring without unit consisting of the finite subsets of $\mathfrak{R}^*$ or of $\mathfrak{R}$. Since the ring $A_{\mathfrak{R}}$ is obtained from the one just described by the

† AH, p. 26, Beispiel 3, and p. 90, footnote.

adjunction of $\mathfrak{R}$ as unit, we conclude that $\mathfrak{E}(A_{\mathfrak{R}})$ is obtained from the space $\mathfrak{R}^*$ by the suitable adjunction of a single point $\mathfrak{p}$ in accordance with Theorem 8. Since the ideal $\mathfrak{a}_{\mathfrak{R}}$ of all nowhere dense sets in $A_{\mathfrak{R}}$ coincides with the class of all finite subsets of $\mathfrak{R}$, we see further that the set $\mathfrak{E}'(\mathfrak{a}_{\mathfrak{R}})$ consists of the point $\mathfrak{p}$ alone. It is now evident that in the map $m(\mathfrak{R}, \mathfrak{E}(A_{\mathfrak{R}}), \Upsilon)$ the sets $\mathfrak{Y}$ are precisely the two-element sets containing the point $\mathfrak{p}$: if $\mathfrak{r}$ is any point of $\mathfrak{R}$ the sets in $A_{\mathfrak{R}}$ which contain it as an interior point are precisely the infinite sets in $A_{\mathfrak{R}}$ which contain it; and the representatives of these sets in $\mathfrak{E}(A_{\mathfrak{R}})$ have as their intersection the set $\mathfrak{Y}(\mathfrak{r})$ consisting of the point $\mathfrak{r}^* = \mathfrak{r}$ in $\mathfrak{R}^* \subset \mathfrak{E}(A_{\mathfrak{R}})$ and the adjoined point $\mathfrak{p}$. If we now apply the construction of Theorem 41 to find the strict T_1-extensions of $\mathfrak{R}$, we find that we must take $Z = X = \Upsilon$: for the only Υ-sets in $\mathfrak{E}(A_{\mathfrak{R}})$ are those which contain the point $\mathfrak{p}$, by virtue of Theorem 34; and every Υ-set therefore either contains a set $\mathfrak{Y}$ or is contained in a set $\mathfrak{Y}$. Thus $\mathfrak{R}$ is absolutely closed with respect to strict T_1-extension, as we wished to prove. To show that $\mathfrak{R}$ is bicompact, we recall that any non-void open set in $\mathfrak{R}$ differs from $\mathfrak{R}$ by a finite set; and it is evident that any family of open sets which covers $\mathfrak{R}$ contains a finite covering subfamily, one member of this subfamily being chosen arbitrarily and the others then being chosen so as to cover the complementary finite set.

If we consider the map described in the preceding proof, we see that it is possible to state the following result.

THEOREM 47. *The space $\mathfrak{R}$ of Theorem 46 has a bicompact strict extension which is a T_0-space absolutely closed with respect to strict T_1-extension.*

Following the construction of Theorem 41, we take $\mathfrak{B} = \mathfrak{E}(A_{\mathfrak{R}})$ and $X = \Upsilon$, and determine Z as the family consisting of the sets in X together with the set $\{\mathfrak{p}\}$. We thus obtain a T_0-space $\mathfrak{Q}$ which is an immediate extension of $\mathfrak{R}$ arising from $\mathfrak{R}$ by the adjunction of a single point ξ corresponding to the set $\{\mathfrak{p}\}$. From the fact that every set in $X = \Upsilon$ contains $\mathfrak{p}$, we see that $\{\xi\}^- = \mathfrak{Q}$. In consequence the only nowhere dense subset of $\mathfrak{R}'$ is the void set. Therefore the complete basic rings $A_{\mathfrak{Q}}$ and $A_{\mathfrak{R}}$ are isomorphic under the correspondence $a \rightarrow a\mathfrak{R}$, as we see by reference to Theorem 39 and the proof of Theorem 41. The map $m(\mathfrak{Q}, \mathfrak{E}(A_{\mathfrak{R}}), Z)$ which was constructed above is therefore equivalent to the complete algebraic map of $\mathfrak{Q}$. As in the preceding theorem, we infer that $\mathfrak{Q}$ is absolutely closed with respect to strict T_1-extension. Moreover, we see that $\mathfrak{Q}$ is not merely an immediate extension of $\mathfrak{R}$ but also a strict extension. Finally we show that $\mathfrak{Q}$ is bicompact. Any family of open sets which covers $\mathfrak{Q}$ intersects $\mathfrak{R}$ in a family of open sets which covers $\mathfrak{R}$; and therefore contains a finite subfamily which covers $\mathfrak{R}$, by virtue of the

bicompactness of $\mathfrak{R}$. To this subfamily, we can adjoin a member of the given family which contains the point ξ, thereby obtaining a finite subfamily which covers the entire space $\mathfrak{Q}$.

THEOREM 48. *In order that a T_0-space $\mathfrak{R}$ be absolutely closed with respect to strict T_1-extension, it is necessary that $\mathfrak{R}$ be bicompact. Any non-bicompact T_0-space $\mathfrak{R}$ becomes, by the suitable adjunction of a single point, a bicompact T_0-space $\mathfrak{Q}$; and $\mathfrak{Q}$ is a strict T_1-extension of $\mathfrak{R}$.*

If $\mathfrak{R}$ is a non-bicompact T_0-space, we define a topology in the class $\mathfrak{Q}=\mathfrak{R}\cup\{\xi\}$ by specifying that the closed subsets of $\mathfrak{Q}$ be (1) the bicompact subspaces of $\mathfrak{R}$ and (2) the sets $\mathfrak{F}\cup\{\xi\}$ where $\mathfrak{F}$ is closed in $\mathfrak{R}$. The indicated subsets clearly have the properties which characterize them as the closed sets under a suitable closure operation.* Since $\{\xi\}^-=\{\xi\}$ on account of the fact that $\{\xi\}=\{\xi\}\cup 0$ is closed in $\mathfrak{Q}$, and since $\mathfrak{R}^-=\mathfrak{Q}$ on account of the fact that $\mathfrak{Q}$ is the only closed subset of $\mathfrak{Q}$ which contains $\mathfrak{R}$, we see that $\mathfrak{Q}$ is an immediate T_1-extension of $\mathfrak{R}$ containing $\mathfrak{R}$ as an open subset. It is easily verified that $\mathfrak{Q}$ is a T_0-space, of course. That $\mathfrak{Q}$ is bicompact, we see as follows: if a family of open sets covers $\mathfrak{Q}$, we choose an arbitrary member $\mathfrak{G}$ such that $\xi\epsilon\mathfrak{G}$, observing that $\mathfrak{G}'$ must be bicompact since it is closed in $\mathfrak{Q}$ and does not contain ξ; and we can then determine a finite covering subfamily by selecting from the given family a finite number of further sets which cover the bicompact set $\mathfrak{G}'$. Finally we show that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. Obviously the nowhere dense subsets of $\mathfrak{R}'$ are 0 and $\{\xi\}$. Thus if $\mathfrak{G}$ is any open set containing ξ, the modified set $\mathfrak{G}-\{\xi\}=\mathfrak{G}\mathfrak{R}$ is open and is contained in $\mathfrak{G}$. If $\mathfrak{G}$ is an open set containing a point $\mathfrak{r}$ in $\mathfrak{R}\subset\mathfrak{Q}$, there exists an open set $\mathfrak{H}$ with the properties $\mathfrak{r}\epsilon\mathfrak{H}\subset\mathfrak{G}$, $\xi\epsilon\mathfrak{H}'$ and the special property that $\mathfrak{H}'\mathfrak{R}$ is closed but not bicompact relative to $\mathfrak{R}$. We find $\mathfrak{H}$ as follows. First, in the space $\mathfrak{R}$, there must exist an open set $\mathfrak{G}_0$ which contains $\mathfrak{r}$ and which has a non-bicompact complement relative to $\mathfrak{R}$: for otherwise $\mathfrak{R}$ would be bicompact, contrary to hypothesis; indeed, any covering family of open sets would contain a set covering $\mathfrak{r}$ and a finite number of further sets covering the bicompact complement of the first. Now $\mathfrak{G}_0$, being open relative to the open subset $\mathfrak{R}$ of $\mathfrak{Q}$, is open in $\mathfrak{Q}$ and does not contain ξ. If we put $\mathfrak{H}=\mathfrak{G}_0\mathfrak{G}$, we see at once that $\mathfrak{q}\epsilon\mathfrak{H}\subset\mathfrak{G}$, $\xi\epsilon\mathfrak{H}'$, $\mathfrak{H}'\mathfrak{R}=\mathfrak{G}'\mathfrak{R}\cup\mathfrak{G}_0'\mathfrak{R}$. Since $\mathfrak{G}_0'\mathfrak{R}$ is not bicompact, $\mathfrak{H}'\mathfrak{R}$ cannot be bicompact. Thus the construction of $\mathfrak{H}$ is completed. We now see that the interior of the modified set $\mathfrak{H}\cup\{\xi\}$ coincides with $\mathfrak{H}$ and is therefore contained in $\mathfrak{G}$: we have $(\mathfrak{H}\cup\{\xi\})'^-=(\mathfrak{H}'\mathfrak{R})^-=\mathfrak{H}'\mathfrak{R}\cup\{\xi\}$ since $\mathfrak{H}'\mathfrak{R}$ is closed in $\mathfrak{R}$ but not bicompact; and thus $(\mathfrak{H}\cup\{\xi\})'^{-\prime}=(\mathfrak{H}'\mathfrak{R}\cup\{\xi\})'$

* See AH, p. 41, Satz VI, and pp. 93–94.

$=(\mathfrak{H} \cup \{\xi\})\mathfrak{R} = \mathfrak{H}$. Consequently, $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$ in accordance with Definition 14.

From the preceding construction, it is apparent that a T_0-space absolutely closed with respect to strict T_1-extension is necessarily bicompact. The converse proposition seems not to be true. Examples might be developed on the basis of Theorem 41 to show that such is the case; but we shall not continue the discussion here.

The study of H-extensions is somewhat easier because of the special conditions noted in Theorem 41. We first have a general characterization.

THEOREM 49. *In order that a T_0-space $\mathfrak{R}$ be closed with respect to immediate or with respect to strict H-extension, it is necessary that in every algebraic map $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ the family $\mathfrak{X}$ cover the closed set $\mathfrak{C}'(\mathfrak{a})$, where $\mathfrak{a}$ is the ideal of nowhere dense sets in the basic ring A defining the map; and it is sufficient that in a single algebraic map the family $\mathfrak{X}$ have this property.*

First, let $\mathfrak{Q}$ be an immediate H-extension of $\mathfrak{R}$; and consider the maps $m(\mathfrak{Q}, \mathfrak{C}(A_{\mathfrak{Q}}), \mathfrak{Z})$, $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{R}}), \mathfrak{Y})$ as described in Theorem 41. Since $\mathfrak{Q}$ is an H-extension of $\mathfrak{R}$, the sets in $\mathfrak{Z}-\mathfrak{X}$, where $\mathfrak{X}$ is the family arising from the map $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{Q}}), \mathfrak{X})$, are disjoint from the sets of $\mathfrak{Z}$. Since any set $\mathfrak{Z}$ is an $\mathfrak{X}$-set, it must have points in common with the set $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}}) \subset \mathfrak{C}(A_{\mathfrak{R}}) \subset \mathfrak{C}(A_{\mathfrak{Q}})$, if we regard $\mathfrak{C}(A_{\mathfrak{R}})$ for convenience as a subset of $\mathfrak{C}(A_{\mathfrak{Q}})$ in accordance with the analysis of Theorem 41. Indeed, we know from Theorem 34 that every set $\mathfrak{Y}$ in $\mathfrak{C}(A_{\mathfrak{R}})$ has a point in common with $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$; and from Theorem 41 that every set $\mathfrak{X}$ is obtained from a corresponding set $\mathfrak{Y}$ by the adjunction of points in $\mathfrak{C}(A_{\mathfrak{Q}})-\mathfrak{C}(A_{\mathfrak{R}})$. Thus, if $\mathfrak{F}$ is any $\mathfrak{X}$-set in $\mathfrak{C}(A_{\mathfrak{Q}})$ and $\mathfrak{G}$ any open set containing $\mathfrak{F}$, then $\mathfrak{G}$ contains some set $\mathfrak{X}$, hence contains a set $\mathfrak{Y}$ which is a subset of $\mathfrak{X}$, and hence contains some point of $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$. It follows that $\mathfrak{F}\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}}) \neq 0$, as we wished to show. Thus, if $\mathfrak{Q}$ is a proper extension of $\mathfrak{R}$, the family $\mathfrak{Z}-\mathfrak{X}$ contains a set $\mathfrak{Z}$ which is disjoint from every set $\mathfrak{X}$ and thus contains a point in $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ which belongs to no set $\mathfrak{Y}$; in other words, $\mathfrak{Y}$ does not cover $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$.

Next, let the map $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{R}}), \mathfrak{Y})$ be such that $\mathfrak{Y}$ does not cover $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$; in particular, let $\mathfrak{p}$ be a point of $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ which belongs to no set $\mathfrak{Y}$. We then carry out the construction of Theorem 41, taking $\mathfrak{B}=\mathfrak{C}(A_{\mathfrak{R}})$, $\mathfrak{X}=\mathfrak{Y}$ and determining $\mathfrak{Z}$ by adjoining $\{\mathfrak{p}\}$ to the family $\mathfrak{X}$. By Theorem 34, $\{\mathfrak{p}\}$ is an $\mathfrak{X}$-set. We therefore obtain an immediate H-extension $\mathfrak{Q}$ of $\mathfrak{R}$ which arises from $\mathfrak{R}$ by the adjunction of a single point ξ corresponding to the set $\{\mathfrak{p}\}$ in $\mathfrak{Z}$. It is evident that $\{\xi\}$ is closed but not open in $\mathfrak{Q}$, and that $\mathfrak{R}$ is open but not closed in $\mathfrak{Q}$. Now we can prove that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. Clearly the nowhere dense subsets of $\mathfrak{R}'$ are 0 and $\{\xi\}$. If $\mathfrak{G}$ is any open set

containing ξ, then the modified set $\mathfrak{G}-\{\xi\}=\mathfrak{G}\mathfrak{R}$ is open and contained in $\mathfrak{G}$. If $\mathfrak{G}$ is any open set containing a point $\mathfrak{r}$ in $\mathfrak{R}$, then the fact that $\mathfrak{r}$ and ξ have the H-separation property implies the existence of an open set $\mathfrak{H}$ such that $\mathfrak{r}\epsilon\mathfrak{H}\subset\mathfrak{G}$, $\xi\epsilon\mathfrak{H}^{-\prime}\subset\mathfrak{H}'$. The modified set $\mathfrak{H}\cup\{\xi\}$ cannot contain ξ as an interior point since, if it did, the set $(\mathfrak{H}\cup\{\xi\})\mathfrak{H}^{-\prime}=\{\xi\}$ would have interior points. Hence the interior of $\mathfrak{H}\cup\{\xi\}$ coincides with $\mathfrak{H}$ and is contained in $\mathfrak{G}$. It follows that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$ in accordance with Definition 14.

If we change our notation slightly to conform with that of Theorem 29, we can therefore assert that the T_0-space $\mathfrak{R}$ is absolutely closed with respect to immediate or strict H-extension if and only if in the complete algebraic map $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{R}}), \mathcal{X})$ the family $\mathcal{X}$ covers the set $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$. The relation between the complete map and an arbitrary algebraic map $m(\mathfrak{R}, \mathfrak{C}(A), \Upsilon)$ described in Theorem 29 now shows that $\mathcal{X}$ covers $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ if and only if Υ covers $\mathfrak{C}'(\mathfrak{a})$ in $\mathfrak{C}(A)$: for the univocal correspondence set up there between $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ and $\mathfrak{C}'(\mathfrak{a})$ carries $\mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ into $\mathfrak{Y}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a})$ and, conversely, carries $\mathfrak{Y}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a})$ back into $\mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$. The present theorem is thus completely proved.

The criterion for absolute closure with respect to H-extension which has just been established can now be replaced by a more familiar criterion, which we give in a somewhat generalized form.

THEOREM 50. *A T_0-space $\mathfrak{R}$ is absolutely closed with respect to immediate or with respect to strict H-extension if and only if every covering family of open sets in $\mathfrak{R}$ contains a finite subfamily of open sets with closures which cover $\mathfrak{R}$. In this criterion, the sufficiency is maintained even if the covering families considered be restricted to be subfamilies of an arbitrary basis for $\mathfrak{R}$.*

We establish this theorem by consideration of the map $m(\mathfrak{R}, \mathfrak{C}(A_{\mathfrak{R}}), \mathcal{X})$, recalling that every open set in $\mathfrak{R}$ is a member of $A_{\mathfrak{R}}$. The open sets in any family $\mathcal{G}$ which covers $\mathfrak{R}$ are represented in $\mathfrak{C}(A_{\mathfrak{R}})$ by sets $\mathfrak{C}(a)$, $a\epsilon A_{\mathfrak{R}}$, with the property that every set $\mathfrak{X}(\mathfrak{r})$ is contained in at least one of them. If $\mathcal{X}$ covers $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$, the family of sets $\mathfrak{C}(a)$ also covers $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$. Thus the bicompactness of $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ establishes the existence of open sets $a_1, \cdots, a_n$ in $\mathcal{G}$ such that $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})\subset\mathfrak{C}(a_1)\cup\cdots\cup\mathfrak{C}(a_n)$. The relation $\mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})\neq 0$ implies the relation $\mathfrak{X}(\mathfrak{r})\mathfrak{C}(a_k)\neq 0$ for at least one index k corresponding to the point $\mathfrak{r}$ in $\mathfrak{R}$. According to Theorem 28, the relations $\mathfrak{X}(\mathfrak{r})\mathfrak{C}(a_k)\neq 0$ and $\mathfrak{r}\epsilon a_k^-$ are equivalent. Hence we obtain the desired relation $a_1^-\vee\cdots\vee a_n^-=\mathfrak{R}$, $a_1\epsilon\mathcal{G}, \cdots, a_n\epsilon\mathcal{G}$. On the other hand, if $\mathcal{X}$ does not cover $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$, let $\mathfrak{p}$ be a point of $\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ which belongs to no set $\mathfrak{X}$. Then if $\mathfrak{r}$ is any point of $\mathfrak{R}$, there exists an element $b(\mathfrak{r})$ in $A_{\mathfrak{R}}$ such that $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{C}(b(\mathfrak{r}))$, $\mathfrak{p}\epsilon\mathfrak{C}'(b(\mathfrak{r}))$. If $a(\mathfrak{r})=b(\mathfrak{r})'^{-\prime}$ is the interior of $b(\mathfrak{r})$, the relation $a(\mathfrak{r})\equiv b(\mathfrak{r})\ (\mathrm{mod}\ \mathfrak{a}_{\mathfrak{R}})$ shows that $\mathfrak{C}(a(\mathfrak{r}))\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})=\mathfrak{C}(b(\mathfrak{r}))\mathfrak{C}'(\mathfrak{a}_{\mathfrak{R}})$ and hence that $\mathfrak{p}\epsilon\mathfrak{C}'(a(\mathfrak{r}))$; and the relation $\mathfrak{r}\epsilon b(\mathfrak{r})'^{-\prime}=a(\mathfrak{r})$ shows in accord-

ance with Theorem 28 that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a(\mathfrak{r}))$. Thus the sets $a(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$, constitute a family $\mathcal{G}$ of open sets covering $\mathfrak{R}$. If $a_1, \cdots, a_n$ is any finite subfamily of $\mathcal{G}$, we have $\mathfrak{p}\epsilon\mathfrak{C}'(a_1) \cdots \mathfrak{C}'(a_n) = \mathfrak{C}(a)$, where $a = a_1' \cdots a_n'$. Since $\{\mathfrak{p}\}$ is an $\mathcal{X}$-set in accordance with Theorem 34, there exists a point $\mathfrak{r}$ such that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a)$. By virtue of Theorem 28, we conclude that $\mathfrak{r}\epsilon a'^{-\prime} = (a_1 \vee \cdots \vee a_n)^{-\prime} = (a_1^- \vee \cdots \vee a_n^-)'$. Hence no finite subfamily of $\mathcal{G}$ has the property that the closures of its members cover $\mathfrak{R}$. On comparing these results with those established in Theorem 49, we see that the first part of the present theorem is proved.

We still wish to show that, if the indicated property holds merely for covering families chosen from an arbitrary fixed basis in $\mathfrak{R}$, then $\mathfrak{R}$ is absolutely closed with respect to immediate or strict H-extension. We obtain this result by showing that we can pass from such restricted covering families to quite general families. Thus let $\mathcal{G}$ be an arbitrary covering family of open sets. If $\mathfrak{r}$ is any point of $\mathfrak{R}$, there exists at least one set $\mathfrak{G}(\mathfrak{r})$ in $\mathcal{G}$ which contains $\mathfrak{r}$; and hence there exists in the given fixed basis a set $\mathfrak{H}(\mathfrak{r})$ such that $\mathfrak{r}\epsilon\mathfrak{H}(\mathfrak{r}) \subset \mathfrak{G}(\mathfrak{r})$. The family $\mathcal{H}$ of all sets $\mathfrak{H}(\mathfrak{r})$ is then a subfamily of the given basis which covers $\mathfrak{R}$. By hypothesis, therefore, there exist sets $\mathfrak{H}_1, \cdots, \mathfrak{H}_n$ in $\mathcal{H}$ such that $\mathfrak{H}_1^- \cup \cdots \cup \mathfrak{H}_n^- = \mathfrak{R}$. The corresponding sets $\mathfrak{G}_1, \cdots, \mathfrak{G}_n$ in $\mathcal{G}$ then have the property $\mathfrak{G}_1^- \cup \cdots \cup \mathfrak{G}_n^- \supset \mathfrak{H}_1^- \cup \cdots \cup \mathfrak{H}_n^- = \mathfrak{R}$. This completes the proof.

Having examined the extension problem in a general way, we shall now proceed to consider specific imbedding and extension theorems, some of which, long formulated, have hitherto remained unproved. Our first result does not concern immediate extensions, but is conveniently stated at this point.

THEOREM 51. *Let $\mathfrak{Q}_\mathfrak{c}$ be the T_0-space obtained by topologizing the family $\mathfrak{Z}$ of all closed sets in the Boolean space $\mathfrak{B}_\mathfrak{c}$ in the usual way, $\mathfrak{c}$ being any infinite cardinal number. Then $\mathfrak{Q}_\mathfrak{c}$ is a universal T_0-space of character $\mathfrak{c}$; in other words, every T_0-space $\mathfrak{R}$ of character not exceeding $\mathfrak{c}$ is topologically equivalent to a subspace of $\mathfrak{Q}_\mathfrak{c}$.*

From Theorem 23, we know that $\mathfrak{Q}_\mathfrak{c}$ is a T_0-space of character not exceeding that of $\mathfrak{B}_\mathfrak{c}$. On the other hand, $\mathfrak{Q}_\mathfrak{c}$ contains $\mathfrak{B}_\mathfrak{c}$ as a subspace, since the one-element subsets of $\mathfrak{B}_\mathfrak{c}$ are members of $\mathfrak{Z}$; and its character is therefore not less than that of $\mathfrak{B}_\mathfrak{c}$. Hence the character of $\mathfrak{Q}_\mathfrak{c}$, like that of $\mathfrak{B}_\mathfrak{c}$, is equal to $\mathfrak{c}$.

Now if $\mathfrak{R}$ is an arbitrary T_0-space of character not exceeding $\mathfrak{c}$, it has a basis of cardinal number not exceeding $\mathfrak{c}$; and the basic ring A generated by such a basis has cardinal number not exceeding $\mathfrak{c}$. We consider the

map $m(\mathfrak{R}, \mathfrak{E}(A), \mathcal{X})$. Since the character of $\mathfrak{E}(A)$ does not exceed $\mathfrak{c}$, we may regard it as a closed subset of the Boolean space $\mathfrak{B}_\mathfrak{c}$ in accordance with the results of Chapter I, §3. Thereby we obtain a map $m(\mathfrak{R}, \mathfrak{B}_\mathfrak{c}, \Upsilon)$ where Υ is the subfamily of $\mathcal{Z}$ consisting of the sets of $\mathcal{X}$, now regarded as closed subsets of $\mathfrak{B}_\mathfrak{c}$. It follows immediately that $\mathcal{Z}$ is an extension of Υ, $\mathfrak{Q}_\mathfrak{c}$ of $\mathfrak{R}$, in accordance with Definition 13.

The discussion of universal finite T_0-spaces is omitted here. Some remarks on this subject will be found in the following section.

THEOREM 52. *Every T_0-space $\mathfrak{R}$ has a strict H-extension $\mathfrak{Q}$ which has the same character as $\mathfrak{R}$ and which is absolutely closed with respect to immediate or with respect to strict H-extension. The space $\mathfrak{R}'$ of points adjoined to $\mathfrak{R}$ in this extension may be taken as a totally-disconnected H-space.**

As in the preceding theorem, we consider a map $m(\mathfrak{R}, \mathfrak{E}(A), \mathcal{X})$ determined by a basic ring A generated by a basis for $\mathfrak{R}$ with minimal cardinal number $\mathfrak{c}$. The character of $\mathfrak{E}(A)$ is then equal to $\mathfrak{c}$, if we exclude the trivial case of finite spaces $\mathfrak{R}$. We obtain the desired space $\mathfrak{Q}$ by adjoining to the family $\mathcal{X}$ all the one-element sets $\{\mathfrak{p}\}$ where $\mathfrak{p}$ is a point of $\mathfrak{E}'(\mathfrak{a})$ which belongs to no set $\mathfrak{X}$. It is then evident that $\mathfrak{Q}$ is an immediate H-extension of $\mathfrak{R}$ in accordance with Theorems 38 and 41. The character of $\mathfrak{Q}$ is not less than that of its subspace $\mathfrak{c}$; and is not greater than $\mathfrak{c}$ because of Theorem 23. Hence the character of $\mathfrak{Q}$, like that of $\mathfrak{R}$, is equal to $\mathfrak{c}$. Moreover, the topology of the set of points adjoined to $\mathfrak{R}$ in the construction of $\mathfrak{Q}$ is equivalent to that of a subset of $\mathfrak{E}'(\mathfrak{a})$; it is therefore a totally-disconnected H-space.

We prove next that $\mathfrak{Q}$ is actually a strict extension of $\mathfrak{R}$. Let us consider an open set specified in $\mathfrak{Q}$ by the relations $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$, $a \in A$, $a \notin \mathfrak{a}$, where $\mathfrak{Z}(\mathfrak{q})$ is either a set in $\mathcal{X}$ or one of the adjoined sets $\{\mathfrak{p}\}$. To this set we adjoin an arbitrary subset of $\mathfrak{R}'$, thus obtaining a set $\mathfrak{F}$ in $\mathfrak{Q}$ and a corresponding family of sets $\mathfrak{Z}(\mathfrak{q})$, $\mathfrak{q} \in \mathfrak{F}$. We determine $\mathfrak{F}'^{-\prime}$, the interior of $\mathfrak{F}$. The sets $\mathfrak{Z}(\mathfrak{q})$ corresponding to $\mathfrak{F}'$ are all the sets $\mathfrak{Z}(\mathfrak{q})$ such that $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(a) \neq 0$, with the exception of the one-element sets which were involved in the adjunction described above. The sets $\mathfrak{Z}(\mathfrak{q})$ corresponding to $\mathfrak{F}'^-$ are precisely the sets $\mathfrak{Z}(\mathfrak{q})$ such that $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(a) \neq 0$, as we see by examining the situation in detail. First, if $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$, the point $\mathfrak{q}$ cannot belong to $\mathfrak{F}'^-$ since the assumed relation means that the open set specified in $\mathfrak{Q}$ by $\mathfrak{Z} \subset \mathfrak{E}(a)$ contains $\mathfrak{q}$ but does not contain any point of $\mathfrak{F}'$. Secondly, if $\mathfrak{Z}(\mathfrak{q}) = \{\mathfrak{p}\}$ is in $\mathfrak{E}'(a)$, then the fact

* The problem of the existence of an immediate H-extension of an H-space $\mathfrak{R}$ under the requirements that the extension shall have the same character as $\mathfrak{R}$ and shall be absolutely closed with respect to immediate H-extension, was proposed by Alexandroff and Urysohn, *Mémoire sur les espaces topologiques compacts*, Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Deel XIV, No. 1 (1929), p. 52.

that $\{\mathfrak{p}\}$ is an $\mathfrak{X}$-set shows that, whenever $\mathfrak{E}(b)$ contains $\mathfrak{p}$, then $\mathfrak{E}(a'b) = \mathfrak{E}'(a)\mathfrak{E}(b)$ contains $\{\mathfrak{p}\}$ and also a set $\mathfrak{X}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$; since the point $\mathfrak{r}$ thus belongs to $\mathfrak{F}'$ in accordance with the fact that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}'(a)$, we conclude that $\mathfrak{q}$ belongs to $\mathfrak{F}'^{-}$. In view of this characterization of the set $\mathfrak{F}'^{-}$, we see that $\mathfrak{F}'^{-\prime}$ consists of those points $\mathfrak{q}$ for which $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$. Now let $\mathfrak{G}$ be any non-void open set in $\mathfrak{Q}$ and let $\mathfrak{q}$ be any point of $\mathfrak{G}$. Since the sets $\mathfrak{G}(a)$ specified in $\mathfrak{Q}$ by the relations $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$, $a\epsilon A$, constitute a basis in $\mathfrak{Q}$, we can find an element a in A and such a corresponding set $\mathfrak{G}(a)$ in $\mathfrak{Q}$ that $\mathfrak{q}\epsilon\mathfrak{G}(a) \subset \mathfrak{G}$. Obviously, the relation $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$ shows that $a\bar{\epsilon}\mathfrak{a}$. Then, if $\mathfrak{H}$ is any subset of $\mathfrak{R}'$, we have $\mathfrak{G}(a)\Delta\mathfrak{H} \subset \mathfrak{F}$, $(\mathfrak{G}(a)\Delta\mathfrak{H})'^{-\prime} \subset \mathfrak{F}'^{-\prime}$, where $\mathfrak{F} = \mathfrak{G}(a) \cup \mathfrak{H}$. As we showed above, $\mathfrak{F}'^{-\prime} = \mathfrak{G}(a)$. Hence we see that $(\mathfrak{G}(a)\Delta\mathfrak{H})'^{-\prime} \subset \mathfrak{G}$. By reference to Definition 14, we conclude that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$.

Finally, we prove that $\mathfrak{Q}$ is closed with respect to immediate or strict H-extension, appealing for this purpose to the criterion stated in Theorem 50. If a family of the basis sets $\mathfrak{G}(a)$, specified by $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(a)$, covers $\mathfrak{Q}$, then the corresponding sets $\mathfrak{E}(a)$ contain every set $\mathfrak{Z}(\mathfrak{q})$ and therefore cover $\mathfrak{E}'(\mathfrak{a})$. Hence there exist elements $a_1, \cdots, a_n$ associated with the given covering family and having the property $\mathfrak{E}'(\mathfrak{a}) \subset \mathfrak{E}(a_1) \cup \cdots \cup \mathfrak{E}(a_n)$. Now we can show that $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a) \neq 0$ implies $\mathfrak{q}\epsilon\mathfrak{G}^{-}(a)$. In order to do so it is sufficient to prove that $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a) \neq 0$ and $\mathfrak{q}\epsilon\mathfrak{G}(b)$ imply $\mathfrak{G}(a)\mathfrak{G}(b) \neq 0$. From the assumed relations $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a) \neq 0$, $\mathfrak{Z}(\mathfrak{q}) \subset \mathfrak{E}(b)$, we obtain the relations $\mathfrak{E}(ab)\mathfrak{E}'(\mathfrak{a}) = \mathfrak{E}(b)\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a) \supset \mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a) \neq 0$ and thus conclude that $ab\bar{\epsilon}\mathfrak{a}$. Hence the interior of ab relative to $\mathfrak{R}$ contains at least one point $\mathfrak{r}$, and $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(ab)$, $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(a)$, $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(b)$, $\mathfrak{r}\epsilon\mathfrak{G}(a)\mathfrak{G}(b)$. If now $\mathfrak{q}$ is any point in $\mathfrak{Q}$, we have $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a_1) \cup \cdots \cup \mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a_n) = \mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})[\mathfrak{E}(a_1) \cup \cdots \cup \mathfrak{E}(a_n)] = \mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a}) \neq 0$, so that $\mathfrak{Z}(\mathfrak{q})\mathfrak{E}'(\mathfrak{a})\mathfrak{E}(a_k) \neq 0$, $\mathfrak{q}\epsilon\mathfrak{G}^{-}(a_k)$ for some index k. Thus we find that $\mathfrak{G}^{-}(a_1) \cup \cdots \cup \mathfrak{G}^{-}(a_n) = \mathfrak{Q}$. By the criterion of Theorem 50, it results that $\mathfrak{Q}$ is absolutely closed with respect to immediate or strict H-extension.

Finally we shall establish an important characterization of the bicompact H-spaces.

THEOREM 53. *In order that an H-space $\mathfrak{R}$ be bicompact, it is necessary and sufficient that every closed subset of $\mathfrak{R}$ be absolutely closed with respect to immediate or with respect to strict H-extension.**

The necessity of the condition is trivial: every closed subset of a bicompact space is bicompact and thus satisfies the criterion given in Theorem 50.

* This theorem was proposed by Alexandroff and Urysohn, *Mémoire sur les espaces topologiques compacts*, Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Deel XIV, No. 1 (1929), p. 50; but was proved only in case $\mathfrak{R}$ is separable.

The sufficiency of the condition is more difficult to establish. In our demonstration we appeal to the following criterion for bicompactness: in order that an H-space be bicompact, it is necessary and sufficient that every transfinite descending sequence of non-void closed subsets have a non-void intersection.* In the given space $\mathfrak{R}$, let $\{\mathfrak{F}_\alpha\}$ be such a sequence, defined for all ordinals α preceding a given ordinal ω: we have $\mathfrak{F}_\alpha \neq 0$, $\mathfrak{F}_\alpha^- = \mathfrak{F}_\alpha$, and $\mathfrak{F}_\alpha \supset \mathfrak{F}_\beta$ for $\alpha < \beta$. We consider the complete algebraic map $m(\mathfrak{R}, \mathfrak{C}(A_\mathfrak{R}), \mathfrak{X})$ and the subfamilies $\mathfrak{X}_\alpha$ of $\mathfrak{X}$ specified by the relation $\mathfrak{r} \epsilon \mathfrak{F}_\alpha$. If $\mathfrak{C}_\alpha(A_\mathfrak{R})$ is the closure of the union of the sets $\mathfrak{X}(\mathfrak{r})$ in $\mathfrak{X}_\alpha$, the map $m(\mathfrak{F}_\alpha, \mathfrak{C}_\alpha(A_\mathfrak{R}), \mathfrak{X}_\alpha)$ then has the properties described in Theorems 36 and 37. In particular the union $\mathfrak{F}(\mathfrak{X}_\alpha)$ of the minimal $\mathfrak{X}_\alpha$-sets is closed and bicompact in $\mathfrak{C}(A_\mathfrak{R})$. If $\alpha_1, \cdots, \alpha_n$ are ordinals preceding ω such that $\alpha_k < \alpha_{k+1}$ for $k = 1, \cdots, n-1$, then the relation $\mathfrak{F}_{\alpha_k} \supset \mathfrak{F}_{\alpha_{k+1}}$ implies $\mathfrak{X}_{\alpha_k} \supset \mathfrak{X}_{\alpha_{k+1}}$. Hence each minimal $\mathfrak{X}_{\alpha_{k+1}}$-set is an $\mathfrak{X}_{\alpha_k}$-set and contains a minimal $\mathfrak{X}_{\alpha_k}$-set, for $k = 1, \cdots, n$. Hence we see that $\mathfrak{F}(\mathfrak{X}_{\alpha_1}) \cdots \mathfrak{F}(\mathfrak{X}_{\alpha_n}) \neq 0$. Since no finite intersection of the closed sets $\mathfrak{F}(\mathfrak{X}_\alpha)$, $\alpha < \omega$, is void, we conclude that there exists at least one point $\mathfrak{p}$ common to them all. By combining results of Theorems 28, 34, 36, 37 and 49, we can now show that this point $\mathfrak{p}$ determines a point $\mathfrak{r}$ in $\mathfrak{R}$ such that $\mathfrak{r} \epsilon \mathfrak{F}_\alpha$ for $\alpha < \omega$. Theorems 36 and 37 permit us to correlate the map $m(\mathfrak{F}_\alpha, \mathfrak{C}_\alpha(A_\mathfrak{R}), \mathfrak{X}_\alpha)$ with an algebraic map $m(\mathfrak{F}_\alpha, \mathfrak{B}_\alpha, \Upsilon_\alpha)$ in such a way that the Υ_α-sets are put in correspondence with the $\mathfrak{X}_\alpha$-sets, inclusion relations being preserved, the sets in Υ_α corresponding biunivocally with the sets in $\mathfrak{X}_\alpha$, and the minimal Υ_α-sets corresponding biunivocally with the minimal $\mathfrak{X}_\alpha$-sets. By Theorem 28, our hypothesis that $\mathfrak{R}$ is an H-space implies that the distinct sets in $\mathfrak{X}_\alpha$ and also those in Υ_α are disjoint. By Theorems 34 and 49, our hypothesis that $\mathfrak{F}_\alpha$ is absolutely closed with respect to immediate or strict extension implies that each minimal Υ_α-set is contained in at least one set belonging to Υ_α. Consequently, we see that each minimal $\mathfrak{X}_\alpha$-set is contained in exactly one set belonging to $\mathfrak{X}_\alpha$. Thus the relation $\mathfrak{p} \epsilon \mathfrak{F}(\mathfrak{X}_\alpha)$ implies the existence of a minimal $\mathfrak{X}_\alpha$-set $\mathfrak{Z}_\alpha$ such that $\mathfrak{p} \epsilon \mathfrak{Z}_\alpha$. This minimal $\mathfrak{X}_\alpha$-set is contained in a set $\mathfrak{X}(\mathfrak{r})$ where $\mathfrak{X}(\mathfrak{r}) \epsilon \mathfrak{X}_\alpha$ or, equivalently, $\mathfrak{r} \epsilon \mathfrak{F}_\alpha$. Since the sets $\mathfrak{X}$ are disjoint, the relation $\mathfrak{p} \epsilon \mathfrak{X}(\mathfrak{r})$ implies that $\mathfrak{r}$ is independent of α for all $\alpha < \omega$. We conclude that $\mathfrak{r}$ is common to the sets $\mathfrak{F}_\alpha$, $\alpha < \omega$. It follows, in accordance with the criterion cited above, that $\mathfrak{R}$ is a bicompact space. The proof of the theorem is thus complete.

5. **Totally-disconnected and discrete spaces.** In concluding the present

* See Alexandroff and Urysohn, Mathematische Annalen, vol. 92 (1924), pp. 258–266; especially pp. 259–261; or Alexandroff and Urysohn, *Mémoire sur les espaces topologiques compacts*, Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Deel XIV, No. 1 (1929), pp. 8–12.

chapter, we shall discuss briefly two special kinds of space: totally-disconnected spaces, as defined in Chapter I, §1, and discrete spaces, as defined below. We first have the following theorem.

THEOREM 54. *A T_0-space $\mathfrak{R}$ is totally-disconnected if and only if it has a biunivocal continuous image $\mathfrak{S}$ which is a subspace of a Boolean space.*

We begin by proving the sufficiency of the stated condition. Let f be a biunivocal continuous correspondence from $\mathfrak{R}$ to $\mathfrak{S}$ where $\mathfrak{S}$ is a subspace of a Boolean space $\mathfrak{B}$. Let $\mathfrak{r}_1$ and $\mathfrak{r}_2$ be distinct points in $\mathfrak{R}$, let $\mathfrak{s}_1=f(\mathfrak{r}_1)$ and $\mathfrak{s}_2=f(\mathfrak{r}_2)$ be their correspondents in $\mathfrak{S}$. By hypothesis $\mathfrak{s}_1\neq\mathfrak{s}_2$, so that there exist closed-and-open sets $\mathfrak{F}_1$ and $\mathfrak{F}_2$ in $\mathfrak{B}$ with the properties $\mathfrak{s}_1\epsilon\mathfrak{F}_1$, $\mathfrak{s}_2\epsilon\mathfrak{F}_2$, $\mathfrak{F}_1\mathfrak{F}_2=0$, $\mathfrak{F}_1 \cup \mathfrak{F}_2=\mathfrak{B}$. The sets $\mathfrak{F}_1\mathfrak{S}$ and $\mathfrak{F}_2\mathfrak{S}$ are closed-and-open relative to $\mathfrak{S}$, so that their antecedents $\mathfrak{H}_1=f^{-1}(\mathfrak{F}_1\mathfrak{S})$, $\mathfrak{H}_2=f^{-1}(\mathfrak{F}_2\mathfrak{S})$ in $\mathfrak{R}$ are closed-and-open. Since $\mathfrak{r}_1\epsilon\mathfrak{H}_1$, $\mathfrak{r}_2\epsilon\mathfrak{H}_2$, $\mathfrak{H}_1\mathfrak{H}_2=0$, and $\mathfrak{H}_1\cup\mathfrak{H}_2=\mathfrak{R}$, we conclude that $\mathfrak{R}$ is totally-disconnected.

We pass now to the necessity of the indicated condition. If $\mathfrak{R}$ is totally-disconnected, the subsets of $\mathfrak{R}$ which are both open and closed constitute a Boolean ring A: for, if a and b are both open and closed, so also are ab, $a \vee b$, a', b', and $a+b=ab'\vee a'b$. The ring A must further be a reduced algebra of classes in $\mathfrak{R}$ in accordance with R Definition 10, since, whenever $\mathfrak{r}_1$ and $\mathfrak{r}_2$ are distinct points in $\mathfrak{R}$, there exists an element a in A such that $\mathfrak{r}_1\epsilon a$, $\mathfrak{r}_2\epsilon a'$. Hence the subclass of A specified by $\mathfrak{r}\,\bar{\epsilon}\, a$ is a prime ideal $\mathfrak{p}(\mathfrak{r})$ in accordance with R Theorem 58. The non-void sets in A can now be used to define a new neighborhood topology over the class $\mathfrak{R}$: to each point of $\mathfrak{R}$ we assign as neighborhoods the members of A which contain it. The resulting topological space, we designate as $\mathfrak{S}$. It is evident that the identical correspondence $\mathfrak{r}\rightarrow\mathfrak{r}$ carries $\mathfrak{R}$ into $\mathfrak{S}$ biunivocally and continuously. Thus, to complete our proof, we have to show that $\mathfrak{S}$ is topologically equivalent to a subspace of a Boolean space. The Boolean ring A is evidently a basic ring for $\mathfrak{S}$. If we construct the algebraic map $m(\mathfrak{S}, \mathfrak{E}(A), \mathcal{X})$, we find that the set $\mathfrak{X}(\mathfrak{r})$ corresponding to a point $\mathfrak{r}$ of $\mathfrak{S}$ consists of a single point in $\mathfrak{E}(A)$, namely, the point $\mathfrak{p}(\mathfrak{r})$ defined above as a prime ideal. For, whenever a is A, we have $a=a^-=a'^{-\prime}$ so that $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{E}(a)$ is equivalent to $\mathfrak{r}\epsilon a$ in accordance with Theorem 28. Thus $\mathfrak{S}$ is topologically equivalent to the subspace of $\mathfrak{E}(A)$ consisting of the points $\mathfrak{p}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{S}$.

We may complete the preceding theorem by the following result.

THEOREM 55. *A T_1-space $\mathfrak{R}$ is topologically equivalent to a subspace of a Boolean space if and only if, whenever $\mathfrak{F}$ is a closed set in $\mathfrak{R}$ and $\mathfrak{r}$ is a point in $\mathfrak{F}'$, there exist closed sets $\mathfrak{F}_1$ and $\mathfrak{F}_2$ in $\mathfrak{R}$ such that $\mathfrak{F}\subset\mathfrak{F}_1$, $\mathfrak{r}\epsilon\mathfrak{F}_2$, $\mathfrak{F}_1\mathfrak{F}_2=0$ and $\mathfrak{F}_1\cup\mathfrak{F}_2=\mathfrak{R}$.*

If $\mathfrak{R}$ is a subspace of a Boolean space $\mathfrak{B}=\mathfrak{E}(A)$, then $\mathfrak{r}\epsilon\mathfrak{F}'$, where $\mathfrak{F}$ is closed in $\mathfrak{R}$, implies the existence of a set $\mathfrak{E}(a)\mathfrak{R}$, where a is in the Boolean ring A, which is open relative to $\mathfrak{R}$ and which has the property that $\mathfrak{F}\mathfrak{E}(a)\mathfrak{R}=0$. Hence we can take $\mathfrak{F}_1=\mathfrak{E}'(a)\mathfrak{R}$, $\mathfrak{F}_2=\mathfrak{E}(a)\mathfrak{R}$.

If $\mathfrak{R}$ has the property indicated in the theorem, the closed-and-open subsets of $\mathfrak{R}$ constitute a basis in $\mathfrak{R}$: for if $\mathfrak{r}$ is any point of $\mathfrak{R}$ and $\mathfrak{G}$ any open set containing $\mathfrak{r}$, the application of this property to $\mathfrak{r}$ and the closed set $\mathfrak{F}=\mathfrak{G}'$ yields a closed-and-open set $\mathfrak{F}_2$ such that $\mathfrak{r}\epsilon\mathfrak{F}_2\subset\mathfrak{G}=\mathfrak{F}'$. The assumption that $\mathfrak{R}$ is a T_1-space implies that any one-element subset of $\mathfrak{R}$ is closed and hence that $\mathfrak{R}$ is totally-disconnected. Accordingly, the associated space $\mathfrak{S}$ constructed from $\mathfrak{R}$ in the manner described in the proof of Theorem 54 is topologically equivalent to $\mathfrak{R}$ and the ring A is a basic ring for $\mathfrak{R}$. We conclude that $\mathfrak{R}$ is equivalent to a subspace of $\mathfrak{E}(A)$.

We pass now to the consideration of discrete spaces, defined as follows:

Definition 18. *A T_0-space $\mathfrak{R}$ is said to be discrete if the closure operation has the property*

$$\Big(\sum_{\mathfrak{A}\epsilon\mathcal{A}}\mathfrak{A}\Big)^- = \sum_{\mathfrak{A}\epsilon\mathcal{A}}\mathfrak{A}^-$$

for any family $\mathcal{A}$ of subsets $\mathfrak{A}$ of $\mathfrak{R}$.

The topology of discrete spaces is described in the following statement, which we give without proof.

Theorem 56. *The following properties of a T_0-space are equivalent:*

(1) *$\mathfrak{R}$ is a discrete space;*

(2) *every union of closed sets is closed;*

(3) *every intersection of open sets is open.*

If the intersection of all open sets containing a fixed point $\mathfrak{r}$ in a discrete T_0-space be denoted by $\mathfrak{G}(\mathfrak{r})$, then

(4) *the sets $\mathfrak{G}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$, constitute a minimal basis for $\mathfrak{R}$—that is, a basis contained in every basis for $\mathfrak{R}$;*

(5) *$\mathfrak{G}(\mathfrak{r}_1)=\mathfrak{G}(\mathfrak{r}_2)$ implies $\mathfrak{r}_1=\mathfrak{r}_2$;*

(6) *if $\mathfrak{G}(\mathfrak{r}_1)\mathfrak{G}(\mathfrak{r}_2)\neq 0$, then $\mathfrak{r}_3\epsilon\mathfrak{G}(\mathfrak{r}_1)\mathfrak{G}(\mathfrak{r}_2)$ implies $\mathfrak{G}(\mathfrak{r}_3)\subset\mathfrak{G}(\mathfrak{r}_1)\mathfrak{G}(\mathfrak{r}_2)$.*

*The relation $\mathfrak{r}_1<\mathfrak{r}_2$ defined by $\mathfrak{r}_1\neq\mathfrak{r}_2$, $\mathfrak{G}(\mathfrak{r}_1)\subset\mathfrak{G}(\mathfrak{r}_2)$ has the properties of a partial order in $\mathfrak{R}$. In terms of this relation, the closure of an arbitrary subset $\mathfrak{A}$ of $\mathfrak{R}$ is the set $\mathfrak{A}^-$ consisting of all points $\mathfrak{r}$ such that $\mathfrak{r}\geqq\mathfrak{r}_0$ for some $\mathfrak{r}_0$ in $\mathfrak{A}$. Every finite T_0-space is discrete.**

We may give a converse of part of this theorem, also without detailed proof.

* See Alexandroff, Comptes Rendus (Paris), vol. 200 (1935), pp. 1649–1651.

THEOREM 57. *If $\mathfrak{R}$ is any partially-ordered class, the introduction of the closure operation defined by*

$$\mathfrak{A}^- = \sum_{\mathfrak{r}_0 \in \mathfrak{A}} \sum_{\mathfrak{r} \geqq \mathfrak{r}_0} \{\mathfrak{r}\}$$

converts $\mathfrak{R}$ into a discrete T_0-space; in this space the ordering relation provided by Theorem 56 *coincides with the one given originally in $\mathfrak{R}$.**

With the aid of the mapping theory, we can now represent a discrete space in terms of Boolean spaces or, equivalently, a partially ordered set in terms of Boolean rings.

THEOREM 58. *Let A be the topological ring for a discrete T_0-space generated by the basis of all sets $\mathfrak{G}(\mathfrak{r})$, described in Theorem* 56. *In the map $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X})$, the relation $\mathfrak{X}(\mathfrak{r}) = \mathfrak{E}(\mathfrak{G}(\mathfrak{r}))$ is valid. Hence the points $\mathfrak{r}$ of $\mathfrak{R}$ and the ordering relation between them are represented by elements of the ring A and the relation of proper inclusion between them. Conversely, any subclass of a Boolean ring A can be regarded as a discrete T_0-space by virtue of the fact that it is partially ordered in terms of the relation of proper inclusion defined in A.*

The theorem is obvious, once we prove that $\mathfrak{X}(\mathfrak{r}) = \mathfrak{E}(\mathfrak{G}(\mathfrak{r}))$. Since $\mathfrak{r}$ is interior to $\mathfrak{G}(\mathfrak{r})$, we have $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(\mathfrak{G}(\mathfrak{r}))$ by virtue of Theorem 28. On the other hand, if $\mathfrak{E}(a) \supset \mathfrak{X}(\mathfrak{r})$ for $a \epsilon A$, we have $\mathfrak{r} \epsilon a'^{-\prime}$ in accordance with Theorem 28; and we conclude that $\mathfrak{G}(\mathfrak{r}) \subset a'^{-\prime} \subset a$, $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{E}(\mathfrak{G}(\mathfrak{r})) \subset \mathfrak{E}(a)$. Hence we must have $\mathfrak{X}(\mathfrak{r}) = \mathfrak{E}(\mathfrak{G}(\mathfrak{r}))$ in accordance with Theorem 26, as we wished to prove. We may note that, by virtue of Theorem 56 (6) the relation $\mathfrak{X}(\mathfrak{r}_1)\mathfrak{X}(\mathfrak{r}_2) \neq 0$ implies the existence of a point $\mathfrak{r}_3$ such that $\mathfrak{X}(\mathfrak{r}_3) \subset \mathfrak{X}(\mathfrak{r}_1)\mathfrak{X}(\mathfrak{r}_2)$: for $\mathfrak{E}(\mathfrak{G}(\mathfrak{r}_1)\mathfrak{G}(\mathfrak{r}_2)) = \mathfrak{E}(\mathfrak{G}(\mathfrak{r}_1))\mathfrak{E}(\mathfrak{G}(\mathfrak{r}_2)) = \mathfrak{X}(\mathfrak{r}_1)\mathfrak{X}(\mathfrak{r}_2) \neq 0$ implies $\mathfrak{G}(\mathfrak{r}_1)\mathfrak{G}(\mathfrak{r}_2) \neq 0$. The interpretation of this theorem which states that any partially ordered set can be imbedded in a Boolean ring with unit has been established abstractly by MacNeille.†

From Theorem 58, we obtain the following special result.

THEOREM 59. *Let $\mathfrak{Q}_n$ be the finite T_0-space obtained by the application of Theorem* 57 *to the partially-ordered class of all non-zero elements in a finite Boolean ring of 2^{n+1} elements, $n = 0, 1, 2, \cdots$. Then $\mathfrak{Q}_n$ is an abstract n-simplex, under a suitable topology. The spaces $\mathfrak{Q}_n$ are universal spaces for all finite T_0-spaces; in other words, a finite T_0-space $\mathfrak{R}$ is topologically equivalent to a subspace of $\mathfrak{Q}_n$ for sufficiently great n.*

* See Alexandroff, Comptes Rendus (Paris), vol. 200 (1935), pp. 1649–1651.

† H. M. MacNeille, *The theory of partially ordered sets*, Harvard doctoral dissertation (1935), not yet published. A summary is given in Proceedings of the National Academy of Sciences, vol. 22 (1936), pp. 45–50.

We can exclude the trivial case where $\mathfrak{R}$ is void. If $\mathfrak{R}$ is a finite non-void T_0-space, it is discrete and can be studied by means of Theorem 58. The ring A evidently is finite, the number of its elements being 2^{n+1}, $n \geqq 0$; and the Boolean space $\mathfrak{C}(A)$ consists of $n+1$ isolated points. The family Z of all sets $\mathfrak{C}(a)$, where $a \epsilon A$ and $a \neq 0$, evidently consists of all non-void subsets of $\mathfrak{C}(A)$. If Z is topologized in the usual way, the resulting T_0-space is equivalent to the space $\mathfrak{Q}_n$: for it is evidently equivalent to the space obtained by regarding the non-zero elements of the ring A as a partially ordered set and introducing the topology described in Theorem 57. In the map $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$, the family $\mathfrak{X}$ is contained in Z. Hence $\mathfrak{R}$ is topologically equivalent to a subspace of $\mathfrak{Q}_n$, as we wished to prove.

As we have just seen, the space $\mathfrak{Q}_n$ is equivalent to the T_0-space constructed by consideration of all the non-void subsets of a Boolean space consisting of $n+1$ isolated points. If we term these subsets "cells," referring to a subset with $k+1$ elements, $k \geqq 0$, in particular as a "k-cell," we can easily make the desired connection with an n-simplex. If one cell is contained in a second, we call it an "edge" of the latter. It is now clear that, as our terminology suggests, the "cells" and the relation of being an "edge" can be exactly represented by the k-dimensional "edges" of an n-dimensional Euclidean cell, for $k=0, 1, \cdots, n$, and the relation of inclusion between them. The topology of the abstract "cell" space, as we have already seen, can be described by the statement that the closure of any given set of "cells" is the class of all "cells" having at least one of the given "cells" as an "edge." Thus we may regard the T_0-space $\mathfrak{Q}_n$ as an abstract n-simplex, the topology of the latter being described in the manner just indicated.

By dualization, it is possible to pass to the more familiar abstract n-simplex in which the closure of any given set of "cells" is the class of all "cells" which are "edges" of at least one of the given "cells." We shall not consider this question further.*

Theorem 59 shows that the finite T_0-spaces are the subspaces of abstract n-simplexes, that is, are abstract complexes. In addition, it completes the results obtained in Theorem 51 for infinite T_0-spaces. From the present point of view, therefore, the spaces $\mathfrak{Q}_c$ may be regarded as abstract simplexes of infinite dimensionality.

Chapter III. Stronger separation conditions

1. **Semi-regular spaces.** In the general theory of Boolean maps as developed in Chapter II, we have considered T_0-, T_1-, and H-spaces on an equal

* For further analysis, see AH, pp. 132–133.

footing. Our results show that the imposition of the T_1- or the H-separation property yields no essential simplification of the theory; in other words, the Boolean maps of T_1-spaces and H-spaces are quite as complicated as those of T_0-spaces. The search for types of space which may admit simplification of the general mapping theory directs our attention to the various stronger separation properties, such as regularity and normality. Upon examination, we discover that the general theory is already in such a form as to suggest the procedure for its own simplification: it is obvious that one should examine first of all the conditions under which the ideal $\mathfrak{a}$ of nowhere dense sets in a basic ring A for a space $\mathfrak{R}$ determines a set of redundancy $\mathfrak{E}(\mathfrak{a})$ in the algebraic map $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X})$. The investigation of this aspect of the mapping theory leads to the introduction of a new type of topological space which proves to occupy a place intermediate between those already considered and the regular spaces. It is these new spaces which we propose to investigate systematically in the present section. We pass later to the study of more special types of space, including the regular spaces.

The formal definition of the spaces to be considered reads as follows:

Definition 19. *A T_0-space $\mathfrak{R}$ is said to be a semi-regular or to be an SR-space if the regular open sets constitute a basis for $\mathfrak{R}$.*

The essential reason that the SR-spaces lead to simplification of the mapping theory is that the nowhere dense sets play a less important rôle than they do in other spaces. This fact is already emphasized in Theorems 24 and 25, and is brought out yet more clearly in the following algebraic theorem.

Theorem 60. *If A is a basic ring in an SR-space $\mathfrak{R}$, if $\mathfrak{a}$ is the ideal of all nowhere dense sets in A, and if $\mathfrak{a}_{\mathfrak{R}}$ is the class of all nowhere dense sets, then there exists a basic ring A^* which is an isomorph of $A/\mathfrak{a}$ and a homomorph of A, the homomorphism $A \to A^*$ being defined by a correspondence $a \to a^*$ such that $a^{-\prime-\prime} < a^* < a^{\prime-\prime-}$, $a \equiv a^*$ (mod $\mathfrak{a}_{\mathfrak{R}}$).*

This theorem reposes essentially upon results given elsewhere by v. Neumann and the writer.† If B is the basic ring generated by A and $\mathfrak{a}_{\mathfrak{R}}$, it is easily seen that the quotient-rings $A/\mathfrak{a}$ and $B/\mathfrak{a}_{\mathfrak{R}}$ are isomorphic: for each residual class (mod $\mathfrak{a}_{\mathfrak{R}}$) in B intersects A in a corresponding residual class (mod $\mathfrak{a}$). The construction of the desired ring A^* is therefore equivalent to the selection, from each residual class (mod $\mathfrak{a}_{\mathfrak{R}}$) in B, of a representative element a^* in such a manner that the representative elements constitute a subring

† v. Neumann and Stone, Fundamenta Mathematicae, vol. 25 (1935), pp. 353–378, especially Theorem 18.

of B. If a is any element of A the correspondence $a \to a^*$, where a^* is the representative of that residual class (mod $\mathfrak{a}_{\mathfrak{R}}$) in B which contains a, obviously defines a homomorphism $A \to A^*$; and it is evident that A^* and $A/\mathfrak{a}$ are isomorphic. Thus we have to solve the $(B, \mathfrak{a}_{\mathfrak{R}})$ representation problem, which was studied in general terms in the paper cited. A solution can be obtained by application of Theorem 18 of that paper. First, let us consider the functions $F(b) = b'^{-\prime-}$, $G(b) = b^{-\prime-\prime}$ defined for all elements b in B. They are connected by the relation $F(b) = G'(b')$. From Theorems 24 and 25 we find that $G(b)$ has the following properties: $G(b) \epsilon B$, $G(b) \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $G(b) = G(c)$ if and only if $b \equiv c$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $G(bc) = G(b)G(c)$. Accordingly, $F(b)$ has the properties: $F(b) \epsilon B$, $F(b) \equiv b$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $F(b) = F(c)$ if and only if $b \equiv c$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $F(b \vee c) = F(b) \vee F(c)$. Secondly, let us consider arbitrary non-void subclasses $\mathfrak{L}$ and $\mathfrak{D}$ in $\mathfrak{a}_{\mathfrak{R}}$ with the property that $c < d$ or, equivalently, $c'd = 0$ for all c in $\mathfrak{L}$ and all d in $\mathfrak{D}$. If we denote by a_0 the union of all sets c in $\mathfrak{L}$, it is evident that $c < a_0 < d$ or, equivalently, $c'a_0 = a_0'd = 0$ for all c in $\mathfrak{L}$ and all d in $\mathfrak{D}$. Since $a_0 < d$ for every d in the non-void class $\mathfrak{D}$ and since $\mathfrak{D}$ contains only nowhere dense sets, a_0 is also nowhere dense, belonging therefore to $\mathfrak{a}_{\mathfrak{R}}$ and to B. Thus the hypotheses of the theorem on which we wish to rely are seen to be satisfied, and the $(B, \mathfrak{a}_{\mathfrak{R}})$ representation problem has a solution. Furthermore, the actual construction of this solution shows that the representative a^* assigned to the residual class (mod $\mathfrak{a}_{\mathfrak{R}}$) containing a given element a in A satisfies the relations $F'(a') < a^* < F(a)$ or, equivalently, $a^{-\prime-\prime} < a^* < a'^{-\prime-}$. The relation $a^* \equiv a$ (mod $\mathfrak{a}_{\mathfrak{R}}$) is evident. We may note that the elements 0 and $e = \mathfrak{R}$ in A have the correspondents $0^* = 0$, $e^* = e$ respectively, by virtue of these relations.

It remains for us to prove that A^* is a basic ring for $\mathfrak{R}$, in other words, that the interiors of the sets in A^* constitute a basis for $\mathfrak{R}$. If $\mathfrak{G}$ is any open set in the SR-space $\mathfrak{R}$ and $\mathfrak{r}$ any point of $\mathfrak{G}$, then there exists a regular open set $\mathfrak{H}$ such that $\mathfrak{r} \epsilon \mathfrak{H} \subset \mathfrak{G}$. Since A is a basic ring for $\mathfrak{R}$, it contains an element a such that $\mathfrak{r} \epsilon a'^{-\prime} \subset \mathfrak{H}$. We shall show that the corresponding element a^* satisfies the relations $\mathfrak{r} \epsilon (a^*)'^{-\prime} \subset \mathfrak{G}$. Using the relations $a < a^-$, $a^{-\prime-\prime} < a^*$, we see that $a'^{-\prime} < (a^-)'^{-\prime} < a^*$, $a'^{-\prime} = (a'^{-\prime})'^{-\prime} < (a^*)'^{-\prime}$, and $\mathfrak{r} \epsilon (a^*)'^{-\prime}$. Using similarly the relation $a^* < a'^{-\prime-}$, we see that $(a^*)'^{-\prime} < (a'^{-\prime-})'^{-\prime} = (a'^{-\prime})^{-\prime-\prime} \subset \mathfrak{H}^{-\prime-\prime} = \mathfrak{H} \subset \mathfrak{G}$. Hence $\mathfrak{r} \epsilon (a^*)'^{-\prime} \subset \mathfrak{G}$, and A^* is a basic ring for $\mathfrak{R}$.

We are now prepared to study the Boolean maps of SR-spaces, beginning with the following fundamental result.

THEOREM 61. *If $\mathfrak{R}$ is an SR-space, if A is a basic ring for $\mathfrak{R}$, if $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ where $\mathfrak{B} = \mathfrak{E}(A)$ is the algebraic map defined by A, and if $\mathfrak{a}$ is the ideal of all nowhere dense sets in A, then $\mathfrak{E}(\mathfrak{a})$ is a set of redundancy for the given map.*

If A^* is the basic ring constructed in Theorem 60, *the suppression of* $\mathfrak{E}(\mathfrak{a})$ *from* $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ *yields a map equivalent to* $m(\mathfrak{R}, \mathfrak{E}(A^*), \mathfrak{X}^*)$; *the latter map has the property that the family* $\mathfrak{X}^*$ *is densely distributed in* $\mathfrak{E}(A^*)$.

The proof that $\mathfrak{E}(\mathfrak{a})$ is a set of redundancy rests upon Theorem 18 above and is similar to the proof of a like result in Theorem 40. In fact, if we put $\mathfrak{Q}=\mathfrak{R}$ and replace the ideal $\mathfrak{b}$ by the ideal $\mathfrak{a}$ in Theorem 40, we see that $\mathfrak{E}(\mathfrak{a})$ is a set of redundancy if $\mathfrak{R}$ has the following property: (P) if $\mathfrak{G}$ is any open set in $\mathfrak{R}$ and $\mathfrak{r}$ any point in $\mathfrak{G}$, then there exists an open set $\mathfrak{H}$ such that $\mathfrak{r}\epsilon\mathfrak{H}\subset\mathfrak{G}$ and such that any set differing from $\mathfrak{H}$ by a nowhere dense set has interior contained in $\mathfrak{G}$. Evidently, an *SR*-space has the property (P): we first choose $\mathfrak{H}$ as a regular open set such that $\mathfrak{r}\epsilon\mathfrak{H}\subset\mathfrak{G}$; and, if $\mathfrak{F}$ is any nowhere dense set, we use the relations $\mathfrak{H}\Delta\mathfrak{F}\subset\mathfrak{H}\cup\mathfrak{F}\subset(\mathfrak{H}\cup\mathfrak{F})^-$ and $\mathfrak{H}\cup\mathfrak{F}\equiv\mathfrak{H}$ (mod $\mathfrak{a}_{\mathfrak{R}}$) to show that $(\mathfrak{H}\Delta\mathfrak{F})'^{-\prime}\subset(\mathfrak{H}\cup\mathfrak{F})'^{-\prime}\subset(\mathfrak{H}\cup\mathfrak{F})^{-\prime-\prime}=\mathfrak{H}^{-\prime-\prime}=\mathfrak{H}\subset\mathfrak{G}$ in accordance with Theorem 25. It follows that $\mathfrak{E}(\mathfrak{a})$ is a set of redundancy for the given map.

The proof that the suppression of $\mathfrak{E}(\mathfrak{a})$ yields a map equivalent to $m(\mathfrak{R}, \mathfrak{E}(A^*), \mathfrak{X}^*)$ is also similar to part of the proof of Theorem 40. Since $A\rightarrow A/\mathfrak{a}\longleftrightarrow A^*$, Theorem 4 shows that the Boolean spaces $\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{E}(A^*)$ are topologically equivalent in such a way that $a\rightarrow a^*$ implies the correspondence of the sets $\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{E}(a^*)$ under this equivalence. If $\mathfrak{r}$ is any point in $\mathfrak{R}$ and if $a\rightarrow a^*$, we shall show that the relations $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})\subset\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{E}(a^*)$ are equivalent; and we can then infer that the topological equivalence between $\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{E}(A^*)$ places $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{X}^*(\mathfrak{r})$ in correspondence. First, let $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})\subset\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$. Then A contains an element b such that $b\equiv a$ (mod $\mathfrak{a}$), $\mathfrak{r}\epsilon b'^{-\prime}$. By Theorems 24 and 25, we must have $b^{-\prime-\prime}=a^{-\prime-\prime}$. The elements a and b have a common correspondent a^* under the homomorphism $A\rightarrow A^*$. In view of the relation $a^{-\prime-\prime}<a^*$, we have $\mathfrak{r}\epsilon b'^{-\prime}<b^{-\prime-\prime}=a^{-\prime-\prime}<a^*$, $\mathfrak{r}\epsilon(a^*)'^{-\prime}$, and $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{E}(a^*)$. On the other hand, let $a\rightarrow a^*$, $\mathfrak{X}^*(\mathfrak{r})\subset\mathfrak{E}(a^*)$. If we use the relation $a'^{-\prime}\equiv a$ (mod $\mathfrak{a}_{\mathfrak{R}}$) in conjunction with Theorems 24 and 25, we find that $(a'^{-\prime-})'^{-\prime}=(a'^{-\prime})^{-\prime-\prime}=a^{-\prime-\prime}$. Since $a^*<a'^{-\prime-}$, we see that $\mathfrak{r}\epsilon(a^*)'^{-\prime}<(a'^{-\prime-})'^{-\prime}=a^{-\prime-\prime}$ or, equivalently, $\mathfrak{r}\notin((a')'^{-\prime})^-$. The basic ring A therefore contains an element b such that $\mathfrak{r}\epsilon b'^{-\prime}$ and $(a')'^{-\prime}b'^{-\prime}=0$. Using Theorem 24, we infer from the relations $a'\equiv(a')'^{-\prime}$ (mod $\mathfrak{a}_{\mathfrak{R}}$), $b\equiv b'^{-\prime}$ (mod $\mathfrak{a}_{\mathfrak{R}}$) that $a'b\equiv 0$ (mod $\mathfrak{a}_{\mathfrak{R}}$). Since $a'b$ is an element of A, the last relation assumes the form $a'b\equiv 0$ (mod $\mathfrak{a}$). Hence we have $\mathfrak{E}(a'b)\subset\mathfrak{E}(\mathfrak{a})$ or, equivalently, $\mathfrak{E}(b)\subset\mathfrak{E}(\mathfrak{a})\cup\mathfrak{E}(a)$. Since $\mathfrak{r}\epsilon b'^{-\prime}$ implies $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{E}(b)\subset\mathfrak{E}(\mathfrak{a})\cup\mathfrak{E}(a)$, we conclude that $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})\subset\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$. In this way we complete the proof that the suppression of $\mathfrak{E}(\mathfrak{a})$ yields a map equivalent to $m(\mathfrak{R}, \mathfrak{E}(A^*), \mathfrak{X}^*)$.

Finally we show that, in the terminology of Definition 4, the family $\mathfrak{X}^*$ is densely distributed in $\mathfrak{C}(A^*)$. Since the ideal $\mathfrak{a}^*$ of nowhere dense sets in A^* consists of the void set alone, we have $\mathfrak{C}'(\mathfrak{a}^*)=\mathfrak{C}(A^*)$. If $\mathfrak{p}^*$ is any point in $\mathfrak{C}(A^*)$, Theorem 34 therefore shows that $\{\mathfrak{p}^*\}$ is an $\mathfrak{X}^*$-set. Consequently, if $\mathfrak{G}^*$ is any non-void open subset of $\mathfrak{C}(A^*)$, it contains not only a set $\{\mathfrak{p}^*\}$ but also a set $\mathfrak{x}^*$ belonging to the family $\mathfrak{X}^*$. This is what we wished to prove.

We now establish a converse of Theorem 61.

THEOREM 62. *If $\mathfrak{R}$ is a T_0-space and if $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ is any Boolean map with the property that Υ is densely distributed in $\mathfrak{B}$, then $\mathfrak{R}$ is an SR-space; and the given map is equivalent to an algebraic map $m(\mathfrak{R}, \mathfrak{C}(A^*), \mathfrak{X}^*)$ of the kind described in Theorem* 61.

We may regard $\mathfrak{B}$ as the representative of an abstract Boolean ring B, and may even take $\mathfrak{B}=\mathfrak{C}(B)$ without loss of generality. The relations $\mathfrak{Y}(\mathfrak{r})\subset\mathfrak{C}(b)$, $b\epsilon B$, define an open subset $\mathfrak{G}(b)$ in $\mathfrak{R}$. We shall show that $\mathfrak{G}(b)$ is even a regular open set. To this end, we prove first that $\mathfrak{G}^{-\prime}(b)=\mathfrak{G}(b')$. If $\mathfrak{r}\epsilon\mathfrak{G}^{-\prime}(b)$, there exists an element c in B such that $\mathfrak{r}\epsilon\mathfrak{G}(c)$ and $\mathfrak{G}(b)\mathfrak{G}(c)=0$: for the sets $\mathfrak{G}(b)$ constitute a basis in $\mathfrak{R}$. We know that $\mathfrak{G}(bc)=\mathfrak{G}(b)\mathfrak{G}(c)=0$. If $bc\neq0$, the fact that Υ is densely distributed in $\mathfrak{C}(B)$ would imply that the open set $\mathfrak{C}(bc)$, being non-void, contains some set $\mathfrak{Y}$; and we could then infer that $\mathfrak{G}(bc)\neq0$. Thus $\mathfrak{G}(bc)=0$ implies $bc=0$. Hence we see that $\mathfrak{Y}(\mathfrak{r})\subset\mathfrak{C}(c)\subset\mathfrak{C}(b')$, $\mathfrak{r}\epsilon\mathfrak{G}(b')$. On the other hand, if $\mathfrak{r}\epsilon\mathfrak{G}(b')$, the relations $\mathfrak{G}(b)\mathfrak{G}(b')=\mathfrak{G}(bb')=\mathfrak{G}(0)=0$ imply $\mathfrak{r}\epsilon\mathfrak{G}^{-\prime}(b)$. Thus the equation $\mathfrak{G}^{-\prime}(b)=\mathfrak{G}(b')$ is valid. It follows immediately that $\mathfrak{G}(b)$ is a regular open set in accordance with the equations $\mathfrak{G}^{-\prime-\prime}(b)=\mathfrak{G}^{-\prime}(b')=\mathfrak{G}(b'')=\mathfrak{G}(b)$. Since the sets $\mathfrak{G}(b)$, $b\epsilon B$, constitute a basis for $\mathfrak{R}$, this space must be an SR-space by Definition 19.

The Boolean ring A generated by the sets $\mathfrak{G}(b)$, $b\epsilon B$, is obviously a basic ring for $\mathfrak{R}$. If $\mathfrak{a}$ is the ideal of nowhere dense sets in A, we can show that $A/\mathfrak{a}$ is isomorphic to the abstract ring B. The proof is based upon the relations $\mathfrak{G}(bc)\equiv\mathfrak{G}(b)\mathfrak{G}(c)$ (mod $\mathfrak{a}$), $\mathfrak{G}(b')\equiv\mathfrak{G}'(b)$ (mod $\mathfrak{a}$), $\mathfrak{G}(b\vee c)\equiv\mathfrak{G}(b)\cup\mathfrak{G}(c)$ (mod $\mathfrak{a}$), and $\mathfrak{G}(b+c)\equiv\mathfrak{G}(b)\Delta\mathfrak{G}(c)$ (mod $\mathfrak{a}$). Since $\mathfrak{G}(bc)=\mathfrak{G}(b)\mathfrak{G}(c)$, the first of these congruences is trivial. The second is established as follows: Theorem 24 shows that $\mathfrak{G}'(b)\equiv(\mathfrak{G}'(b))'^{-\prime}$ (mod $\mathfrak{a}_{\mathfrak{R}}$) and the result noted above shows that $\mathfrak{G}^{-\prime}(b)\equiv\mathfrak{G}(b')$ (mod $\mathfrak{a}_{\mathfrak{R}}$); combining these congruences we see that $\mathfrak{G}'(b)\equiv\mathfrak{G}(b')$ (mod $\mathfrak{a}_{\mathfrak{R}}$); and, observing that $\mathfrak{G}'(b)$ and $\mathfrak{G}(b')$ are both in A, we can rewrite the last congruence in the form $\mathfrak{G}(b')\equiv\mathfrak{G}'(b)$ (mod $\mathfrak{a}$). The remaining congruences then follow algebraically from the first two. Any element a in A is expressible as a polynomial in terms of the sets $\mathfrak{G}(b)$, $b\epsilon B$, and the operations $\cdot$ and Δ, since the unit e in B has the property that $\mathfrak{G}(e)=\mathfrak{R}$.

Consequently the congruences established above enable us to write $a \equiv \mathfrak{G}(b) \pmod{\mathfrak{a}_{\mathfrak{R}}}$ for a suitable element b in B. Moreover, this congruence becomes $a \equiv \mathfrak{G}(b) \pmod{\mathfrak{a}}$ by virtue of the fact that a and $\mathfrak{G}(b)$ both belong to A. In this congruence the element b is uniquely determined: for $\mathfrak{G}(b) \equiv \mathfrak{G}(c) \pmod{\mathfrak{a}}$ implies $\mathfrak{G}(b+c) \equiv 0 \pmod{\mathfrak{a}}$; the fact that $\mathfrak{G}(b+c)$ is a regular open set then implies $\mathfrak{G}(b+c)=0$; and it follows, as we observed above, that $b+c=0$ or, equivalently, $b=c$. The correspondence from A to B defined by putting $a \rightarrow b$ whenever $a \equiv \mathfrak{G}(b) \pmod{\mathfrak{a}}$ is now seen to be a homomorphism: it is univocal; and the correspondences $a_1 \rightarrow b_1$, $a_2 \rightarrow b_2$ evidently imply $a_1a_2 \equiv \mathfrak{G}(b_1b_2) \pmod{\mathfrak{a}}$, $a_1+a_2 \equiv \mathfrak{G}(b_1+b_2) \pmod{\mathfrak{a}}$, or $a_1a_2 \rightarrow b_1b_2$, $a_1+a_2 \rightarrow b_1+b_2$. Since the relations $a \rightarrow 0$ and $a \equiv \mathfrak{G}(0) \equiv 0 \pmod{\mathfrak{a}}$ are equivalent, the ideal defined in A by the homomorphism $A \rightarrow B$ coincides with $\mathfrak{a}$; and the rings $A/\mathfrak{a}$ and B are therefore isomorphic.

By applying Theorem 36 we shall next prove that the map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ is equivalent to that obtained from $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X})$ by the suppression of the set $\mathfrak{E}(\mathfrak{a})$; and Theorem 61 then justifies the assertion that $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ is equivalent to an algebraic map $m(\mathfrak{R}, \mathfrak{E}(A^*), \mathfrak{X}^*)$. First we recall that the isomorphism $A/\mathfrak{a} \longleftrightarrow B$ defines a topological equivalence between $\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{B} = \mathfrak{E}(B)$ which places the sets $\mathfrak{E}(a)\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{E}(b)$ in correspondence whenever the homomorphism $A \rightarrow B$ carries a into b. Thus the topological equivalence may be described in the following explicit terms: if $\mathfrak{p}$ is any point in $\mathfrak{E}'(\mathfrak{a})$, its image in $\mathfrak{E}(B)$ is the unique point $\mathfrak{q}$ which is common to all sets $\mathfrak{E}(b)$ where $\mathfrak{p} \in \mathfrak{E}(a)$ and $a \rightarrow b$; and, conversely, if $\mathfrak{q}$ is any point in $\mathfrak{E}(B)$, its image in $\mathfrak{E}'(\mathfrak{a})$ is the unique point $\mathfrak{p}$ common to all sets $\mathfrak{E}(a)$ where $a \rightarrow b$ and $\mathfrak{q} \in \mathfrak{E}(b)$. We now compare this correspondence between $\mathfrak{B} = \mathfrak{E}(B)$ and $\mathfrak{E}'(\mathfrak{a})$ with the correspondence from Υ-sets to $\mathfrak{X}$-sets constructed in Theorem 36. By Theorem 34, the minimal $\mathfrak{X}$-sets in $\mathfrak{E}(A)$ are precisely the one-element subsets of $\mathfrak{E}'(\mathfrak{a})$; and, by virtue of the fact that Υ is densely distributed in $\mathfrak{B}$, the minimal Υ-sets in $\mathfrak{B} = \mathfrak{E}(B)$ are precisely the one-element subsets of $\mathfrak{E}(B)$. From the preceding description of the topological equivalence between $\mathfrak{E}'(\mathfrak{a})$ and $\mathfrak{E}(B)$, and from the information obtained in Theorem 36, we therefore see that the two correspondences under discussion have the same effect upon the minimal sets: the topological equivalence places the minimal $\mathfrak{X}$-sets in biunivocal correspondence with the minimal Υ-sets in a way which coincides with that described in Theorem 36. Since the correspondence of Theorem 36 carries $\mathfrak{Y}(\mathfrak{r})$ into $\mathfrak{X}(\mathfrak{r})$ and preserves inclusion-relations, we infer that the topological equivalence places $\mathfrak{Y}(\mathfrak{r})$ in correspondence with $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})$: for the minimal Υ-sets contained in $\mathfrak{Y}(\mathfrak{r})$ are in correspondence with the minimal $\mathfrak{X}$-sets contained in $\mathfrak{X}(\mathfrak{r})$; and these two families of minimal sets have $\mathfrak{Y}(\mathfrak{r})$ and $\mathfrak{X}(\mathfrak{r})\mathfrak{E}'(\mathfrak{a})$ as their respective unions. We have thereby proved that

$m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ is equivalent to the map obtained from $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ by the suppression of the set $\mathfrak{C}(\mathfrak{a})$.

It is useful to summarize the results of Theorems 61 and 62 in the following terms:

THEOREM 63. *The following properties of a T_0-space $\mathfrak{R}$ are equivalent:*

(1) *$\mathfrak{R}$ is an SR-space;*

(2) *$\mathfrak{R}$ has a Boolean map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$, where Υ is densely distributed in $\mathfrak{B}$;*

(3) *$\mathfrak{R}$ has an algebraic map $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ in which $\mathfrak{C}(\mathfrak{a})$ is a set of redundancy.*

If a space $\mathfrak{R}$ has any of these equivalent properties, then those of its Boolean maps $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ which have the property of being densely distributed in $\mathfrak{B}$ are characterized as the irredundant algebraic maps of $\mathfrak{R}$; each such map can be constructed from a suitable algebraic map $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ by suppression of the corresponding set $\mathfrak{C}(\mathfrak{a})$; and this construction, applied to an arbitrary algebraic map, yields a map of the indicated type.

With Theorem 63, the theory of maps for semi-regular spaces is brought to a satisfactory conclusion. That this development of the theory corresponds to a real rather than an apparent specialization is shown by the following result.

THEOREM 64. *There exist T_0-, T_1- and H-spaces which are not SR-spaces.*

The T_0-space discussed in Theorem 47 is neither a T_1-space nor an SR-space: we have already seen that it is not a T_1-space; and by reference to Theorem 61 we now see that it cannot be an SR-space. Similarly, the T_1-space discussed in Theorem 46 is neither an H-space nor an SR-space. A space described in AH provides an example of an H-space which is not an SR-space.* The points of this space $\mathfrak{R}$ are the real numbers r, $0 \leqq r \leqq 1$. A neighborhood system is introduced in $\mathfrak{R}$ as follows: if $0 < r_0 < 1$, the neighborhoods of r_0 are the sets $a < r < b$ where $0 < a < r_0 < b < 1$; if $r = 1$, the neighborhoods of r are the sets $a < r \leqq 1$ where $0 < a$; and if $r_0 = 0$, the neighborhoods of r_0 are the sets $\mathfrak{G}(a)$ consisting of all points r such that $0 \leqq r < a$, $r \neq 1/n$ where $a < 1$ and $n = 2, 3, 4, \cdots$. It is easily verified that $\mathfrak{R}$ is an H-space. If $\mathfrak{G}$ is a regular open set containing the point 0, then $\mathfrak{G}$ contains some neighborhood $\mathfrak{G}(a)$ of 0. Since $\mathfrak{G}^-(a)$ obviously consists of all points r such that $0 \leqq r \leqq a$, we see that $\mathfrak{G}^{-\prime-\prime}(a)$ consists of all points r such that $0 \leqq r < a$. The relations $\mathfrak{G}(a) \subset \mathfrak{G}$, $\mathfrak{G}^-(a) \subset \mathfrak{G}^-$ show that $\mathfrak{G} = \mathfrak{G}^{-\prime-\prime} \supset \mathfrak{G}^{-\prime-\prime}(a)$. Hence $\mathfrak{G}$ is contained in no neighborhood of the point 0. Thus $\mathfrak{R}$ is not an SR-space. We can show in addition that $\mathfrak{R}$ is absolutely closed with respect

* AH, p. 31, Beispiel 1.

to immediate or strict H-extension. Indeed, if an infinite family of distinct open sets covers $\mathfrak{R}$, at least one of them contains a neighborhood $\mathfrak{G}(a)$ of 0; and the remaining sets cover the set $a \leqq r \leqq 1$. Since the topology of the subspace $a \leqq r \leqq 1$ is the usual topology, we see that the given family contains sets $\mathfrak{G}_1, \cdots, \mathfrak{G}_n$, where $\mathfrak{G}_1 \supset \mathfrak{G}(a)$ and $\mathfrak{G}_2, \cdots, \mathfrak{G}_n$ cover the set $a \leqq r \leqq 1$. Thus the sets $\mathfrak{G}_1^-, \cdots, \mathfrak{G}_n^-$ cover $\mathfrak{R}$; and $\mathfrak{R}$ is absolutely closed with respect to immediate or strict H-extension in accordance with Theorem 50.

The simplification of the mapping theory in the case of SR-spaces leads to a corresponding simplification of the theory of extensions. The chief result assumes the following form.

THEOREM 65. *If a T_0-space $\mathfrak{R}$ has an immediate extension $\mathfrak{Q}$ which is an SR-space, then $\mathfrak{R}$ is itself an SR-space and $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. All SR-spaces $\mathfrak{Q}$ which are strict extensions of an SR-space $\mathfrak{R}$ can be obtained by the following construction: in a Boolean map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ where Υ is densely distributed in $\mathfrak{B}$, the family Υ is augmented by the adjunction of Υ-sets to provide a family Z; and the familiar topology is then imposed upon Z. Conversely, this construction always yields an SR-space $\mathfrak{Q}$ which is a strict extension of $\mathfrak{R}$. The space $\mathfrak{Q}$ is a T_1-extension of $\mathfrak{R}$ if and only if no set in $Z-\Upsilon$ contains or is contained in any distinct set in Z; and $\mathfrak{Q}$ is an H-extension of $\mathfrak{R}$ if and only if no set in $Z-\Upsilon$ has points in common with any distinct set in Z.*

If $\mathfrak{R}$ has an immediate extension $\mathfrak{Q}$ which is an SR-space, we consider the maps described in Theorem 40. If $\mathfrak{a}$ is the ideal of all nowhere dense sets in A, then the suppression of the set $\mathfrak{C}(\mathfrak{a})$ from $m(\mathfrak{Q}, \mathfrak{C}(A), Z)$ yields a map $m(\mathfrak{Q}, \mathfrak{C}'(\mathfrak{a}), Z^*)$ in accordance with Theorem 61. If we consider only those members of the families Z and Z^* which represent points of the subspace $\mathfrak{R}$, we obtain maps $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ and $m(\mathfrak{R}, \mathfrak{C}'(\mathfrak{a}), \mathfrak{X}^*)$. It is evident that the second of these maps can be constructed from the first by the suppression of the set $\mathfrak{C}(\mathfrak{a})$. We now show that $\mathfrak{X}^*$ is densely distributed in $\mathfrak{C}'(\mathfrak{a})$. If $\mathfrak{p}$ is any point in $\mathfrak{C}'(\mathfrak{a})$, then $\{\mathfrak{p}\}$ is a Z-set in accordance with Theorem 34; but Theorem 40 shows that $\{\mathfrak{p}\}$ is also an $\mathfrak{X}$-set. Hence $\mathfrak{p}\epsilon\mathfrak{C}(a)$ implies the existence of a set $\mathfrak{X}(\mathfrak{r})$ in $\mathfrak{X}$ such that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a)$. Consequently $\mathfrak{p}\epsilon\mathfrak{C}'(\mathfrak{a})\mathfrak{C}(a)$ implies the existence of a corresponding set $\mathfrak{X}^*(\mathfrak{r}) = \mathfrak{X}(\mathfrak{r})\mathfrak{C}'(\mathfrak{a})$ in $\mathfrak{X}^*$ such that $\mathfrak{X}^*(\mathfrak{r}) \subset \mathfrak{C}'(\mathfrak{a})\mathfrak{C}(a)$, as we wished to prove. Since $\mathfrak{X}^*$ is densely distributed in $\mathfrak{C}'(\mathfrak{a})$, the existence of the map $m(\mathfrak{R}, \mathfrak{C}'(\mathfrak{a}), \mathfrak{X}^*)$ implies that $\mathfrak{R}$ is an SR-space in accordance with Theorem 62. Furthermore the given extension $\mathfrak{Q}$ can evidently be obtained from the map $m(\mathfrak{R}, \mathfrak{C}'(\mathfrak{a}), \mathfrak{X}^*)$ by the construction described in the statement of the theorem. We can easily verify that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$ in accordance with Definition 14: if $\mathfrak{q}$ is any point of the SR-space $\mathfrak{Q}$ and $\mathfrak{G}$ any open set containing $\mathfrak{q}$, there exists a regular open

set $\mathfrak{H}$ such that $\mathfrak{q}\epsilon\mathfrak{H}\subset\mathfrak{G}$; and if $\mathfrak{F}$ is any nowhere dense set, not merely one which is contained in $\mathfrak{R}'$, we have $(\mathfrak{H}\Delta\mathfrak{F})'^{-\prime}\subset(\mathfrak{H}\Delta\mathfrak{F})^{-\prime-\prime}=\mathfrak{H}^{-\prime-\prime}=\mathfrak{H}\subset\mathfrak{G}$.

The construction described in the statement of the theorem yields an immediate extension $\mathfrak{Q}$ of $\mathfrak{R}$ in accordance with Theorem 38. Since the family Υ is assumed to be densely distributed in $\mathfrak{B}$, the family Z is also densely distributed in $\mathfrak{B}$; and the space $\mathfrak{Q}$ must therefore be an *SR*-space in accordance with Theorem 62. It follows that $\mathfrak{Q}$ is a strict extension of $\mathfrak{R}$. The condition that $\mathfrak{Q}$ be a T_1-extension of $\mathfrak{R}$ has already been established in Theorem 38. So also has the sufficiency of the condition that $\mathfrak{Q}$ be an H-extension of $\mathfrak{R}$. We shall now prove that this condition is necessary: if $\mathfrak{Z}(\mathfrak{q})\epsilon Z-X$ and $\mathfrak{Z}(\mathfrak{r})\epsilon Z$ have the property $\mathfrak{Z}(\mathfrak{q})\mathfrak{Z}(\mathfrak{r})\neq 0$, then any two open sets $\mathfrak{G}(\mathfrak{q})$ and $\mathfrak{G}(\mathfrak{r})$ in $\mathfrak{B}$ with the properties $\mathfrak{Z}(\mathfrak{q})\subset\mathfrak{G}(\mathfrak{q})$, $\mathfrak{Z}(\mathfrak{r})\subset\mathfrak{G}(\mathfrak{r})$ necessarily have the property $\mathfrak{G}(\mathfrak{q})\mathfrak{G}(\mathfrak{r})\neq 0$; and since Z is densely distributed in $\mathfrak{B}$, there exists a set $\mathfrak{Z}$ in Z such that $\mathfrak{Z}\subset\mathfrak{G}(\mathfrak{q})\mathfrak{G}(\mathfrak{r})$. Accordingly the open sets $\mathfrak{H}(\mathfrak{q})$, $\mathfrak{H}(\mathfrak{r})$ specified in $\mathfrak{Q}$ by the respective relations $\mathfrak{Z}\subset\mathfrak{G}(\mathfrak{q})$, $\mathfrak{Z}\subset\mathfrak{G}(\mathfrak{r})$ have the properties $\mathfrak{q}\epsilon\mathfrak{H}(\mathfrak{q})$, $\mathfrak{r}\epsilon\mathfrak{H}(\mathfrak{r})$, $\mathfrak{H}(\mathfrak{q})\mathfrak{H}(\mathfrak{r})\neq 0$, where $\mathfrak{q}\epsilon\mathfrak{R}'$ and $\mathfrak{r}\epsilon\mathfrak{Q}$. Since the sets $\mathfrak{H}$ specified by $\mathfrak{Z}\subset\mathfrak{G}$ constitute a basis for $\mathfrak{Q}$, the points $\mathfrak{q}$ and $\mathfrak{r}$ obviously do not have the H-separation property; and $\mathfrak{Q}$ is not an H-extension of $\mathfrak{R}$. This completes the proof.

It is also easy to specialize Theorem 52 for the case of *SR*-spaces. We obtain the following result.

THEOREM 66. *Every SR-space $\mathfrak{R}$ of infinite character $\mathfrak{c}$ has a strict H-extension $\mathfrak{Q}$ which has character $\mathfrak{c}$, which is absolutely closed with respect to immediate or strict H-extension, and which is an SR-space. The points adjoined to $\mathfrak{R}$ in this extension may be assumed to constitute a totally-disconnected H-space.*

We first construct a Boolean map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ where $\mathfrak{B}$ has character $\mathfrak{c}$ and Υ is densely distributed in $\mathfrak{B}$. We start this construction by selecting a basis of cardinal number $\mathfrak{c}$ for $\mathfrak{R}$. The basic ring A generated by this basis evidently has cardinal number $\mathfrak{c}$; and the representative Boolean space $\mathfrak{E}(A)$ has character $\mathfrak{c}$. The suppression of the set $\mathfrak{E}(\mathfrak{a})$ from $\mathfrak{E}(A)$ therefore leaves a Boolean space $\mathfrak{E}'(\mathfrak{a})$ of character not exceeding $\mathfrak{c}$. Theorem 61 then shows that we can obtain the desired map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ by taking $\mathfrak{B}=\mathfrak{E}'(\mathfrak{a})$. According to Theorem 23, the character of $\mathfrak{R}$ cannot exceed that of $\mathfrak{E}'(\mathfrak{a})$. Hence the character of $\mathfrak{E}'(\mathfrak{a})$ must be equal to $\mathfrak{c}$. Now we apply the construction of Theorem 65 to obtain the desired extension $\mathfrak{Q}$ by means of the map $m(\mathfrak{Q}, \mathfrak{B}, Z)$. The family Z is formed by adjoining to Υ all one-element subsets $\{\mathfrak{p}\}$ in $\mathfrak{B}$, where $\mathfrak{p}$ belongs to none of the sets in Υ. Theorems 62 and 63 show that $m(\mathfrak{Q}, \mathfrak{B}, Z)$ is an algebraic map; by construction, the family Z covers $\mathfrak{B}$; and hence Theorem 49 shows that $\mathfrak{Q}$ is absolutely closed with respect to

immediate or strict H-extension. By the reasoning given in the proof of Theorem 52, the points adjoined to $\mathfrak{R}$ in this extension constitute a totally-disconnected H-space.

In conclusion, we remark upon the connections between the results of the present section and those given in Theorem 29. By comparing Theorems 29, 61, 62, and 63, we are able to state the following result.

THEOREM 67. *If $\mathfrak{R}$ is a T_0-space and $\mathfrak{R}^*$ the associated continuous image constructed in Theorem 29, then $\mathfrak{R}^*$ is an SR-space. If $\mathfrak{R}$ is itself an SR-space, then $\mathfrak{R}$ and $\mathfrak{R}^*$ are topologically equivalent.*

2. **Regular spaces.** We pass now to the consideration of a stronger separation property which has long played an important rôle in general topology. We introduce it through one of the standard definitions.

DEFINITION 20. *A T_0-space is said to be regular or to be an $\mathfrak{R}$-space if, whenever $\mathfrak{G}$ is an open set in $\mathfrak{R}$ and $\mathfrak{r}$ is a point in $\mathfrak{G}$, there exists an open set $\mathfrak{H}$ such that $\mathfrak{r}\epsilon\mathfrak{H}$, $\mathfrak{H}^- \subset \mathfrak{G}$.*

The relative strength of this separation property is disclosed by the following theorem.

THEOREM 68. *An R-space is both an SR-space and an H-space.*

That an R-space is an SR-space is evident: for, if $\mathfrak{H}$ is the open set of Definition 20, the regular open set $\mathfrak{H}^{-\prime-\prime}$ has the properties $\mathfrak{r}\epsilon\mathfrak{H} = \mathfrak{H}^{\prime-\prime} \subset \mathfrak{H}^{-\prime-\prime}$, $\mathfrak{H}^{-\prime-\prime} \subset \mathfrak{G}^{\prime-\prime} = \mathfrak{G}$. In order to prove that an R-space is an H-space, we proceed as follows: if $\mathfrak{p}$ and $\mathfrak{q}$ are distinct points in $\mathfrak{R}$, the assumption that $\mathfrak{R}$ is a T_0-space implies that, the notation being properly chosen, $\mathfrak{p}$ is contained in an open set to which $\mathfrak{q}$ does not belong or, equivalently, that $\mathfrak{p}\epsilon\{\mathfrak{q}\}^{-\prime}$; hence there exists an open set $\mathfrak{H}$ such that $\mathfrak{p}\epsilon\mathfrak{H}$, $\mathfrak{H}^- \subset \{\mathfrak{q}\}^{-\prime}$; and the points $\mathfrak{p}$ and $\mathfrak{q}$ therefore have the H-separation property in accordance with the relations $\mathfrak{p}\epsilon\mathfrak{H}$, $\mathfrak{q}\epsilon\mathfrak{H}^{-\prime}$, $\mathfrak{H}\mathfrak{H}^{-\prime} = 0$.

In view of Theorem 68, the simplified form of the mapping theory developed for SR-spaces applies in particular to R-spaces. It is possible, however, to carry the study of R-spaces somewhat further than this.

THEOREM 69. *If a T_0-space $\mathfrak{R}$ has a continuous Boolean map $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$, then $\mathfrak{R}$ is an R-space. Conversely, if $\mathfrak{R}$ is an R-space, its irredundant algebraic maps $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$, and likewise its general algebraic maps $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$ where A contains a basis for $\mathfrak{R}$, are all continuous maps.*

Let $m(\mathfrak{R}, \mathfrak{B}, \Upsilon)$ be a continuous Boolean map of a T_0-space $\mathfrak{R}$. If $\mathfrak{G}$ is an open set in $\mathfrak{R}$ and $\mathfrak{r}$ is any point in $\mathfrak{G}$, then there exists a set $\mathfrak{G}_1$ which has the properties $\mathfrak{r}\epsilon\mathfrak{G}_1 \subset \mathfrak{G}$ and which is specified by a bicompact open set $\mathfrak{F}_1$ in $\mathfrak{B}$ through the relations $\mathfrak{Y} \subset \mathfrak{F}_1$, $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{F}_1$: for the sets so specified constitute a

basis for $\mathfrak{R}$ in accordance with Theorem 23. Since the given map is continuous, there exists an open set $\mathfrak{F}_0$, which we may assume also to be bicompact, such that $\mathfrak{Y}(\mathfrak{r}) \subset \mathfrak{F}_0$ and such that $\mathfrak{Y}\mathfrak{F}_0 \neq 0$ implies $\mathfrak{Y} \subset \mathfrak{F}_1$. We denote by $\mathfrak{H}$ the open set in $\mathfrak{R}$ specified by the relation $\mathfrak{Y} \subset \mathfrak{F}_0$. It is obvious that $\mathfrak{r} \epsilon \mathfrak{H} \subset \mathfrak{G}_1 \subset \mathfrak{G}$. The set $\mathfrak{F}_0'$ is also a bicompact open set in $\mathfrak{B}$. The open set specified in $\mathfrak{R}$ by the relation $\mathfrak{Y} \subset \mathfrak{F}_0'$ has as its complement a closed set $\mathfrak{F}$ which is described by the relation $\mathfrak{Y}\mathfrak{F}_0 \neq 0$. Hence we have $\mathfrak{H} \subset \mathfrak{F} \subset \mathfrak{G}_1 \subset \mathfrak{G}$. Since $\mathfrak{H}^- \subset \mathfrak{F}^- = \mathfrak{F}$, the open set $\mathfrak{H}$ satisfies the relations $\mathfrak{r} \epsilon \mathfrak{H}$, $\mathfrak{H}^- \subset \mathfrak{G}$. The space $\mathfrak{R}$ is therefore an R-space.

If $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ is an algebraic map of an R-space $\mathfrak{R}$, let $\mathfrak{X}(\mathfrak{r})$ be an arbitrary member of the family $\mathcal{X}$ and let $\mathfrak{G}$ be an open subset of $\mathfrak{C}(A)$ which contains $\mathfrak{X}(\mathfrak{r})$. Then there exists an element a in A such that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(a) \subset \mathfrak{G}$, $\mathfrak{r} \epsilon a'^{-\prime}$. By hypothesis $\mathfrak{R}$ contains an open set $\mathfrak{H}$ such that $\mathfrak{r} \epsilon \mathfrak{H}$, $\mathfrak{H}^- \subset a'^{-\prime}$. The basic ring A then contains an element b such that $\mathfrak{r} \epsilon b'^{-\prime} \subset \mathfrak{H}$. If we could use the equation $b'^{-\prime -} = b^-$, we could conclude that $b^- = b'^{-\prime -} \subset \mathfrak{H}^- \subset a'^{-\prime}$; and, by virtue of Theorem 28, that $\mathfrak{X}\mathfrak{C}(b) \neq 0$ implies $\mathfrak{X} \subset \mathfrak{C}(a)$. Since $\mathfrak{r} \epsilon b'^{-\prime}$ implies $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{C}(b)$ we could then show that the element $c = ab$ has the following properties: $\mathfrak{X}(\mathfrak{r})$ is contained in $\mathfrak{C}(c) = \mathfrak{C}(a)\mathfrak{C}(b) \subset \mathfrak{G}$; and $\mathfrak{X}\mathfrak{C}(c) \neq 0$ implies $\mathfrak{X}\mathfrak{C}(b) \neq 0$ and hence $\mathfrak{X} \subset \mathfrak{C}(a) \subset \mathfrak{G}$. The family $\mathcal{X}$ would thus be continuous in accordance with Definition 5, the map $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ continuous in accordance with Definition 7. We therefore consider the conditions under which we can assert that $b'^{-\prime -} = b^-$. If A contains a basis for $\mathfrak{R}$, we may take b as an open set; and we then have $b'^{-\prime} = b$, $b'^{-\prime -} = b^-$. Again, if the map $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ is irredundant, Theorems 63 and 68 show that $\mathcal{X}$ is densely distributed; and the proof of Theorem 62 then shows that the equation $b'^{-\prime} = \mathfrak{G}(b)$, which is valid because both members are specified by the relation $\mathfrak{X} \subset \mathfrak{C}(b)$, implies $b'^{-\prime -} = \mathfrak{G}^-(b) = (\mathfrak{G}^{-\prime}(b))' = \mathfrak{G}'(b') = ((b')'^{-\prime})' = b^-$. Thus the map $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ is continuous under either of the conditions stated in the theorem.

By combining Theorem 69 with Theorem 22, we obtain the following result.

Theorem 70. *In order that the algebraic map* $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ *define a continuous univocal correspondence* $\mathfrak{X}(\mathfrak{R}) \rightarrow \mathfrak{r}$ *from the subspace* $\mathfrak{S}(\mathcal{X}) = \sum_{\mathfrak{r}\epsilon\mathfrak{R}} \mathfrak{X}(\mathfrak{r})$ *of* $\mathfrak{C}(A)$ *to the* T_0*-space* $\mathfrak{R}$, *it is necessary and sufficient that* $\mathfrak{R}$ *be an R-space and* $m(\mathfrak{R}, \mathfrak{C}(A), \mathcal{X})$ *a continuous map. All R-spaces are thus found among the continuous images of subspaces of bicompact Boolean spaces.*

It is now possible to recover from the mapping theory a well-known elementary criterion for bicompactness.*

* AH, pp. 91–92.

THEOREM 71. *In order that an H-space be bicompact, it is necessary and sufficient that it be an R-space absolutely closed with respect to immediate or strict H-extension.*

We have already noted in Theorem 53 that a bicompact H-space is absolutely closed with respect to immediate or strict H-extension. We can now show that a bicompact H-space $\mathfrak{R}$ is an R-space. If $\mathfrak{G}$ is any open set in $\mathfrak{R}$ and $\mathfrak{r}$ is any point in $\mathfrak{G}$, there exist open sets $\mathfrak{G}(\mathfrak{p})$, $\mathfrak{H}(\mathfrak{p})$ corresponding to an arbitrary point $\mathfrak{p}$ in $\mathfrak{G}'$ such that $\mathfrak{p}\epsilon\mathfrak{G}(\mathfrak{p})$, $\mathfrak{r}\epsilon\mathfrak{H}(\mathfrak{p})$, $\mathfrak{G}(\mathfrak{p})\mathfrak{H}(\mathfrak{p})=0$. Since $\mathfrak{G}'$ is closed, it is bicompact; and there exist points $\mathfrak{p}_1, \cdots, \mathfrak{p}_n$ in $\mathfrak{G}'$ such that $\mathfrak{G}(\mathfrak{p}_1) \cup \cdots \cup \mathfrak{G}(\mathfrak{p}_n) \supset \mathfrak{G}'$, $\mathfrak{r}\epsilon\mathfrak{H}(\mathfrak{p}_1) \cdots \mathfrak{H}(\mathfrak{p}_n) \subset \mathfrak{G}'(\mathfrak{p}_1) \cdots \mathfrak{G}'(\mathfrak{p}_n) \subset \mathfrak{G}$. If we define $\mathfrak{H}$ as the open set $\mathfrak{H}=\mathfrak{H}(\mathfrak{p}_1) \cdots \mathfrak{H}(\mathfrak{p}_n)$, we find that $\mathfrak{r}\epsilon\mathfrak{H}$, $\mathfrak{H}^- \subset (\mathfrak{G}'(\mathfrak{p}_1) \cdots \mathfrak{G}'(\mathfrak{p}_n))^- \subset \mathfrak{G}$. On the other hand, if $\mathfrak{R}$ is an R-space absolutely closed with respect to immediate or strict H-extension, we construct an irredundant algebraic map $m(\mathfrak{R}, \mathfrak{E}(A), \mathfrak{X})$ in accordance with Theorem 61. Since $\mathfrak{R}$ is absolutely closed in the indicated sense, the family $\mathfrak{X}$ covers $\mathfrak{E}(A)$ in accordance with Theorem 49. Since $\mathfrak{R}$ is an R-space, Theorems 69 and 70 show that $\mathfrak{R}$ is a continuous image of the space $\mathfrak{E}(A)=\mathfrak{S}(\mathfrak{X})$. It follows from the bicompactness of $\mathfrak{E}(A)$ that $\mathfrak{R}$ is bicompact. In fact, any family of open sets which covers $\mathfrak{R}$ has a family of antecedents which are open and cover $\mathfrak{E}(A)$; and, since a finite number of these antecedents suffices to cover $\mathfrak{E}(A)$, a finite number of sets in the given family likewise suffices to cover $\mathfrak{R}$.

The proof just given for Theorem 71 can evidently be rearranged in such a way as to establish the following result, which we give without further formal discussion.

THEOREM 72. *Among all H-spaces, the bicompact spaces are characterized topologically as the continuous images of bicompact Boolean spaces.**

As a consequence of Theorem 71 we can now show that the SR-separation property is essentially weaker than the R-separation property.

THEOREM 73. *There exists an SR-space which is an H-space but not an R-space. Such a space may be constructed so as to be absolutely closed with respect to immediate or strict H-extension.*

Let us suppose, on the contrary, that every space which is both an SR-space and an H-space is necessarily an R-space. If $\mathfrak{R}$ is an arbitrary R-space, we apply Theorem 66 to construct an immediate H-extension $\mathfrak{Q}$ of $\mathfrak{R}$ which is an SR-space absolutely closed with respect to immediate or strict H-extension. Theorem 42 shows that the extension $\mathfrak{Q}$ so constructed is an

---

* See the material related to this result in AH, pp. 95–98.

H-space. By hypothesis, $\mathfrak{Q}$ is therefore an R-space. Theorem 71 therefore implies that $\mathfrak{Q}$ is bicompact. Thus our initial assumption leads to the following proposition: every R-space can be imbedded as a subspace in a bicompact H-space. Now an example due to Tychonoff* shows that this proposition is false. Whenever $\mathfrak{R}$ is restricted to have character $\aleph_0$, however, the proposition is known to be true. Hence, if we start our construction with the R-space $\mathfrak{R}$ given by Tychonoff, we obtain a space $\mathfrak{Q}$ of the type described in the theorem. The character of the space $\mathfrak{Q}$, being equal to that of $\mathfrak{R}$, must exceed $\aleph_0$.

We can rid ourselves of this restriction on the character of $\mathfrak{Q}$ by following a different process of construction. The Cantor discontinuum $\mathfrak{D}$ is a bicompact Boolean space of character $\aleph_0$. In $\mathfrak{D}$ we introduce a family $\mathcal{X}$ of disjoint closed sets as follows: first we form the two-element sets $\mathfrak{X}_n$ consisting of the points $x=3^{-n}=2\sum_{\alpha=n+1}^{\infty}3^{-\alpha}$, $x=1-3^{-n}=2\sum_{\alpha=1}^{n}3^{-\alpha}$ for $n=1, 2, 3, \cdots$, respectively; and then we form all the one-element sets $\{x\}$ where x belongs to no set $\mathfrak{X}_n$. We see at once that the family $\mathcal{X}$ is densely distributed in $\mathfrak{D}$ but is not continuous. On topologizing the family $\mathcal{X}$ in the usual way, we obtain a space $\mathfrak{Q}$ and a Boolean map $m(\mathfrak{Q}, \mathfrak{D}, \mathcal{X})$. The character of $\mathfrak{Q}$ is $\aleph_0$. Theorem 63 shows that $\mathfrak{Q}$ is an SR-space and that the map $m(\mathfrak{Q}, \mathfrak{D}, \mathcal{X})$ is an irredundant algebraic map. Since $\mathcal{X}$ covers $\mathfrak{D}$, Theorem 49 shows that $\mathfrak{Q}$ is absolutely closed with respect to immediate or strict H-extension. Since the sets in $\mathcal{X}$ are disjoint, the space $\mathfrak{Q}$ is an H-space. Theorem 69 shows finally that $\mathfrak{Q}$ is not an R-space. A similar construction could evidently be carried out in an arbitrary bicompact Boolean space of infinite character.

In conclusion we may pose the problem of discovering what restrictions upon SR-spaces imply regularity. We have seen above that certain obvious restrictions are not sufficient. In this connection, it is natural to consider the following question: what SR-spaces have the property that every subspace is also an SR-space; and, in particular, what H-spaces have this property? From Theorem 65, we know that all everywhere dense subspaces of an SR-space are SR-spaces. The problem therefore reduces to the restricted problem concerning closed subspaces alone.

3. **Completely regular spaces.** As we recalled at the close of the preceding section, there exist regular spaces which cannot be imbedded in bicompact H-spaces. Consequently any separation property which characterizes the subspaces of bicompact H-spaces must be stronger than the property of regularity. Such a stronger property was discovered by Tychonoff, who termed it "complete regularity."* We propose to investigate the influence of

* Tychonoff, Mathematische Annalen, vol. 102 (1930), pp. 544–561.

this property upon the mapping theory. First, let us state the fundamental definition and some of its elementary consequences.

DEFINITION 21. *A T_0-space $\mathfrak{R}$ is said to be completely regular or to be a CR-space if it has any one of the three equivalent properties:*

(1) *if $\mathfrak{F}$ is any closed set in $\mathfrak{R}$ and $\mathfrak{r}_0$ is any point in $\mathfrak{F}'$, then there exists a continuous real function f defined in $\mathfrak{R}$ such that $f(\mathfrak{r}_0)=0$, $f(\mathfrak{r})=1$ when $\mathfrak{r}\epsilon\mathfrak{F}$, $0\leqq f(\mathfrak{r})\leqq 1$;*

(2) *the open sets $\mathfrak{R}(\alpha<f<\beta)$, where α and β are arbitrary real numbers and f is an arbitrary bounded continuous real function in $\mathfrak{R}$, constitute a basis for $\mathfrak{R}$;**

(3) *if $\mathfrak{G}$ is any open set in $\mathfrak{R}$ and $\mathfrak{r}_0$ is any point in $\mathfrak{G}$, then there exists a family of open sets $\mathfrak{G}_\tau$ defined for all rational numbers τ, $0<\tau<1$, such that $\mathfrak{r}_0\epsilon\mathfrak{G}_\tau$, $\mathfrak{G}_\tau^- \subset \mathfrak{G}$, $\sigma<\tau$ implies $\mathfrak{G}_\sigma^- \subset \mathfrak{G}_\tau$.*

The equivalence of these three conditions is proved by well-known arguments, which we sketch briefly. First, (1) implies (2): if $\mathfrak{r}_0\epsilon\mathfrak{G}$, we take the function f corresponding to the closed set $\mathfrak{F}=\mathfrak{G}'$ and the point $\mathfrak{r}_0$ as described in (1) and observe that $\mathfrak{G}_0=\mathfrak{R}(-1/2<f<1/2)$ has the properties $\mathfrak{r}_0\epsilon\mathfrak{G}_0\subset\mathfrak{G}$. Similarly, (1) implies (3): if $\mathfrak{r}_0\epsilon\mathfrak{G}$, we again take f as the function corresponding to $\mathfrak{F}=\mathfrak{G}'$ and $\mathfrak{r}_0$, and put $\mathfrak{G}_\tau=\mathfrak{R}(f<\tau)$, $0<\tau<1$. Next (2) implies (1): if $\mathfrak{r}_0\epsilon\mathfrak{F}'$, there exists an open set $\mathfrak{R}(\alpha<g<\beta)$ as described in (2) such that $\mathfrak{r}_0\epsilon\mathfrak{R}(\alpha<g<\beta)\subset\mathfrak{F}'$; and, if we put $\gamma=\min\,(g(\mathfrak{r}_0)-\alpha,\ \beta-g(\mathfrak{r}_0))\neq 0$, $f(\mathfrak{r})=\min\,(1,\ |g(\mathfrak{r})-g(\mathfrak{r}_0)|/\gamma)$, we find that f has the properties demanded in (1). Finally (3) implies (1). Taking $\mathfrak{G}=\mathfrak{F}'$, we form the family $\mathfrak{G}_\tau$, $0<\tau<1$, corresponding to $\mathfrak{G}$ and $\mathfrak{r}_0$ in accordance with (3); and we define $\mathfrak{G}_\tau=0$ for rational numbers τ, $\tau<0$, and $\mathfrak{G}_\tau=\mathfrak{R}$ for rational numbers τ, $\tau>1$. We then define $f(\mathfrak{r})$ as the greatest lower bound of the numbers τ such that $\mathfrak{r}\epsilon\mathfrak{G}_\tau$. It is evident that $f(\mathfrak{r}_0)=0$, $0\leqq f(\mathfrak{r})\leqq 1$, $f(\mathfrak{r})=1$ in $\mathfrak{G}'=\mathfrak{F}$. To prove that f is continuous, we must show that $\mathfrak{R}(\alpha<f<\beta)$ is an open set. If $\mathfrak{r}$ is any point in this set, there exist rational numbers ρ, σ, τ such that $\alpha<\sigma<\rho<f(\mathfrak{r})<\tau<\beta$. We then have $\mathfrak{r}\epsilon\mathfrak{G}_\rho'\mathfrak{G}_\tau\subset\mathfrak{G}_\sigma^{-\prime}\mathfrak{G}_\tau$. On the other hand, if $\mathfrak{r}\epsilon\mathfrak{G}_\sigma^{-\prime}\mathfrak{G}_\tau$ where $\alpha<\sigma<\tau<\beta$, we see that $\mathfrak{r}\,\bar\epsilon\,\mathfrak{G}_\sigma$, $\mathfrak{r}\epsilon\mathfrak{G}_\tau$, hence that $\sigma\leqq f(\mathfrak{r})\leqq\tau$, and hence that $\alpha<f(\mathfrak{r})<\beta$. Thus $\mathfrak{R}(\alpha<f<\beta)$ is the union of the open sets $\mathfrak{G}_\sigma^{-\prime}\mathfrak{G}_\tau$ where $\alpha<\sigma<\tau<\beta$; and $\mathfrak{R}(\alpha<f<\beta)$ is also open, as we wished to prove.

We remark that the condition (3) can be expressed, without reference to the rational numbers, in terms of order alone. For this purpose, we introduce the concept of strong inclusion: the set $\mathfrak{F}_1$ is said to be strongly included in the set $\mathfrak{F}_2$, if $\mathfrak{F}_1^-\subset\mathfrak{F}_2$. Then we may replace (3) by the equivalent (3′): if $\mathfrak{r}_0\epsilon\mathfrak{G}$,

* By $\mathfrak{R}(P)$ we mean the set of all points in $\mathfrak{R}$ which have the property P, here the property $\alpha<f<\beta$.

there exists a countable family of open sets which contain $\mathfrak{r}_0$, which are contained in $\mathfrak{G}$, and which are simply ordered by the relation of strong inclusion in such a way that there is no first, no last, and no pair without an intermediate third. The property (3′) may therefore be taken as characterizing the completely regular spaces in a purely topological manner. In practice, the properties (1) and (2) are the ones which are technically useful.

The relation of CR-spaces to other types of space is indicated in the following theorem.

THEOREM 74. *A CR-space is an R-space, and hence both an SR- and an H-space; every normal space is a CR-space. Every subspace of a CR-space is a CR-space. There exist R-spaces which are not CR-spaces, and CR-spaces which are not normal spaces.*

From Definition 21 (3), it is evident that every CR-space is an R-space in accordance with Definition 20. Theorem 68 then shows that a CR-space is both an SR-space and an H-space. The proof that every normal space is a CR-space is given by establishing property (3) of Definition 21; the explicit construction of the family $\mathfrak{G}_\tau$ is a familiar one.* By use of Definition 21 (1), it is easily seen that every subspace of a CR-space is also a CR-space. Examples given by Tychonoff justify the final statement of the theorem.†

Since the definition of CR-spaces involves, directly or indirectly, the topological relations between general topological spaces and the real number system, it is clear that we cannot investigate them without injecting new methods into the discussion. In order to continue the algebraic tendencies of the present paper and also in order to obtain the maximum of new information, we propose to study the topological ring of all bounded continuous real functions in an arbitrary T_0-space $\mathfrak{R}$. We shall find that the mapping theory for CR-spaces is closely connected with the theory of the corresponding function-rings. Moreover, we shall find that this connection enables us to give a full account of the problem of imbedding CR-spaces as everywhere dense subspaces in bicompact H-spaces. In the subsequent discussion we shall denote the real number field by the letter R; and we shall not distinguish between this field and its isomorphs.

THEOREM 75. *If $\mathfrak{M}$ is the class of all bounded continuous real functions defined in a T_0-space $\mathfrak{R}$, if the sum and product of such functions are defined in the usual way, and if the norm $\|f\|$ of a function f is defined as the least upper bound of the numbers $|f(\mathfrak{r})|$, $\mathfrak{r}\in\mathfrak{R}$, then $\mathfrak{M}$ is a topological ring in the following sense:*

* See AH, p. 74.

† Tychonoff, Mathematische Annalen, vol. 102 (1930), pp. 544–561.

(1) *under the operations of addition and multiplication, $\mathfrak{M}$ is a commutative ring containing R as a subring;*
(2) *if $\|f-g\|$ is introduced as the distance between f and g, then $\mathfrak{M}$ is a complete metric space;*
(3) *in the metric topology of (2), the ring operations are continuous, the polynomials $f+g$, fg being continuous functions of their arguments f and g.*

If multiplication by f be restricted to the case where $f\epsilon R$, then $\mathfrak{M}$ is a Banach space. The operation of forming the absolute value $|f|$ of a function f in $\mathfrak{M}$ is continuous in the metric topology of $\mathfrak{M}$. Any subring of $\mathfrak{M}$ which contains αf together with f whenever $\alpha\epsilon R$ and which is a closed subset of $\mathfrak{M}$ has the property that it contains $|f|$ together with f.

The verification of the statements made in this theorem may be left to the reader, as the detailed proofs are all familiar from elementary analysis. We may point out that metric convergence of a sequence in $\mathfrak{M}$ is equivalent to uniform pointwise convergence of the function-sequence in $\mathfrak{R}$. The final statement of the theorem involves an application of the Weierstrass approximation theorem: if $\alpha=\|f\|$, there exists a polynomial $p_n(x)$ such that $\big||x|-p_n(x)\big|\leqq 1/n$ for $-\alpha\leqq x\leqq\alpha$ and $p_n(0)=0$; and it follows that $\big\||f|-p_n(f)\big\|\leqq 1/n$, $p_n(f)=\sum_{\nu=1}^{N(n)}\alpha_\nu(n)f^\nu$. For general discussions of the subject matter of the present theorem, the reader is referred to the literature.†

We shall find it convenient to introduce a few special descriptive terms for later use.

Definition 22. *The ring $\mathfrak{M}$ of Theorem 75 is called the function-ring of the space $\mathfrak{R}$. A closed subring of $\mathfrak{M}$ which contains the subfield R is called an analytical subring of $\mathfrak{M}$. If $\mathfrak{N}$ and $\mathfrak{N}^*$ are homomorphic subrings of function-rings $\mathfrak{M}$ and $\mathfrak{M}^*$ respectively with the property that $f\rightarrow f^*$ implies $|f|\rightarrow|f^*|$ and $\|f\|=\|f^*\|$, then the homomorphism $\mathfrak{N}\rightarrow\mathfrak{N}^*$ is necessarily an isomorphism and is called an analytical isomorphism; and $\mathfrak{N}$ and $\mathfrak{N}^*$ are said to be analytically isomorphic.*

There are several elementary theorems which we may mention informally at this point. Thus, we see that the analytical subring generated by a non-void subclass $\mathfrak{M}_0$ of $\mathfrak{M}$ is the closure of the subring generated by R and $\mathfrak{M}_0$. It is obvious from Theorem 75 that an analytical subring contains $|f|$ together with f. If $\mathfrak{N}^*$ is analytically isomorphic to a closed subring $\mathfrak{N}$, then $\mathfrak{N}^*$ is closed. If $\mathfrak{N}^*$ is analytically isomorphic to an analytical subring $\mathfrak{N}$, then there exist an analytical subring $\mathfrak{N}^{**}$ containing $\mathfrak{N}^*$ and a function

† See, for instance, Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, especially pp. 11, 53; Chittenden, these Transactions, vol. 31 (1929), pp. 290–321.

θ^* in $\mathfrak{N}^*$ which assumes only the values 0 and 1 and does not vanish identically, such that $\mathfrak{N}^*$ is the class of all functions $f^*=\theta^* f^{**}$ where $f^{**}\epsilon\mathfrak{N}^{**}$. To prove this assertion, we first consider the image R^* in $\mathfrak{N}^*$ of the real field R in $\mathfrak{N}$. If $\alpha\epsilon R$ and $\alpha\geqq 0$, then $\alpha\rightarrow f^*$ implies $f^*\geqq 0$: for $\alpha=|\alpha|$ implies $f^*=|f^*|$. If θ^* is the correspondent of 1, then $\theta^*=(\theta^*)^2$ so that θ^* assumes only the values 0 and 1. If f is any element in $\mathfrak{N}$, then $f\rightarrow f^*$ implies $1\cdot f\rightarrow\theta^* f^*$ and hence $f^*=\theta^* f^*$. Since $\mathfrak{N}$ and $\mathfrak{N}^*$ are isomorphic it is therefore evident that θ^* does not vanish identically. If α is the rational number m/n and $\alpha\rightarrow f^*$, the relation $n\cdot\alpha=m\cdot 1$ or $\alpha+\cdots+\alpha=1+\cdots+1$ implies $nf^*=m\theta^*$, and hence $f^*=m/n\theta^*=\alpha\theta^*$. If α is arbitrary and α_1 and α_2 are rational approximants to α satisfying the inequalities $\alpha_1\leqq\alpha\leqq\alpha_2$, then $\alpha\rightarrow f^*$ implies $\alpha_1\theta^*\leqq f^*\leqq\alpha_2\theta^*$; and we can therefore conclude that $f^*=\alpha\theta^*$. We now form the class $\mathfrak{N}^{**}$ of all functions $f^{**}=f^*+\alpha(1-\theta^*)$ where $f^*\epsilon\mathfrak{N}^*$ and $\alpha\epsilon R$. Since $\mathfrak{N}^*$ is closed and since $f^*=\theta^* f^*$ is in $\mathfrak{N}^*$, it is evident that $\mathfrak{N}^{**}$ is a closed subring of $\mathfrak{M}^*$. Moreover, $\mathfrak{N}^{**}$ contains R: for, if α is any real number, $\alpha\theta^*$ is in $\mathfrak{N}^*$ and $\alpha=\theta^*+\alpha(1-\theta^*)$ is in $\mathfrak{N}^{**}$. Thus $\mathfrak{N}^{**}$ is an analytical subring of $\mathfrak{M}^*$; and $\mathfrak{N}^*$ is obtained from $\mathfrak{N}^{**}$ in the manner described above. In the space $\mathfrak{R}^*$ in which $\mathfrak{M}^*$ is defined, the equation $\theta^*=1$ defines a closed subspace $\mathfrak{R}^{**}$. Obviously, $\mathfrak{N}^*$ is an analytical subring of the function-ring for $\mathfrak{R}^{**}$. In order that θ^* be identically equal to 1 or, equivalently, that $\mathfrak{N}^*$ be an analytical subring of $\mathfrak{M}^*$, it is necessary and sufficient that $\mathfrak{N}^*$ contain 1; and sufficient that $\mathfrak{R}^*$ be connected.

We now establish a fundamental theorem concerning ideals in function-rings.

THEOREM 76. *A necessary and sufficient condition that an ideal $\mathfrak{A}$ in the function-ring $\mathfrak{M}$ be divisorless is that the quotient-ring $\mathfrak{M}/\mathfrak{A}$ be isomorphic to R. Such an ideal is necessarily closed and prime in $\mathfrak{M}$; and the homomorphism $\mathfrak{M}\rightarrow R$ determined by it has the property that $f\rightarrow\alpha$ implies $|f|\rightarrow|\alpha|$ and* g.l.b. $f\leqq\alpha\leqq$ l.u.b. f.

If $\mathfrak{M}/\mathfrak{A}$ is isomorphic to R, the fact that R is a field shows that R has no ideals other than R and $\{0\}$; and it follows that $\mathfrak{A}$ has no ideal divisors other than $\mathfrak{M}$ and $\mathfrak{A}$, in other words, that $\mathfrak{A}$ is divisorless. Since the quotient-ring is a field, $\mathfrak{A}$ is also prime. It is evident that $\mathfrak{A}$ does not contain the unit 1 in $\mathfrak{M}$.

If $\mathfrak{A}$ is a divisorless ideal in $\mathfrak{M}$, the inequality $\mathfrak{A}\neq\mathfrak{M}$ shows that 1 does not belong to $\mathfrak{A}$. Since the closure $\mathfrak{A}^-$ is also an ideal, the relation $\mathfrak{A}\subset\mathfrak{A}^-$ implies that $\mathfrak{A}=\mathfrak{A}^-$ or $\mathfrak{M}=\mathfrak{A}^-$. The equation $\mathfrak{M}=\mathfrak{A}^-$ holds if and only if 1 is in $\mathfrak{A}^-$. Now if 1 were an element of $\mathfrak{A}^-$, there would exist in $\mathfrak{A}$ a uniformly convergent sequence of functions with 1 as limit. Hence $\mathfrak{A}$ would contain a function f such that g.l.b. $f>0$. Since we would then have $1/f\epsilon\mathfrak{M}$ and $f\epsilon\mathfrak{A}$, we

could conclude that $1=f\cdot(1/f)\epsilon\mathfrak{A}$, contrary to fact. The equation $\mathfrak{A}=\mathfrak{A}^-$ is therefore valid; and $\mathfrak{A}$ is closed in $\mathfrak{M}$. We can next show that $\mathfrak{M}/\mathfrak{A}$ is a field. If f is any element in $\mathfrak{M}$, then the ideal generated by f and $\mathfrak{A}$ coincides with $\mathfrak{M}$. Hence there exist functions g and h in $\mathfrak{M}$ and in $\mathfrak{A}$ respectively such that $fg+h=1$. Thus the relation $f\not\equiv 0 \pmod{\mathfrak{A}}$ implies the existence of a function g in $\mathfrak{M}$ such that $fg\equiv 1 \pmod{\mathfrak{A}}$, as we wished to prove. It is clear that $\mathfrak{A}$ contains no member of R other than 0; and it follows that two members of R are congruent $\pmod{\mathfrak{A}}$ if and only if they are identical. Hence the field $\mathfrak{M}/\mathfrak{A}$ contains R as a subfield.

In order to prove that $\mathfrak{M}/\mathfrak{A}$ coincides with its subfield R, we shall apply the theory of ordered fields.* The first step is to classify the elements of $\mathfrak{M}/\mathfrak{A}$ as positive, negative, or zero in such a way that $f+g$ and fg are positive, $-f$ negative, when f and g are positive. In this classification we wish to maintain the natural classification of the subfield R. For convenience we shall define a similar classification of $\mathfrak{M}$ in such a way that functions congruent $\pmod{\mathfrak{A}}$ are assigned to the same class. We say that f is positive if $f\equiv|f| \pmod{\mathfrak{A}}$, $f\not\equiv 0 \pmod{\mathfrak{A}}$, and that f is negative if $-f$ is positive; and denote the class of positive elements in $\mathfrak{M}$ by $\mathfrak{P}$, the class of negative elements by $\mathfrak{N}$. In order to justify this classification, we prove the following propositions: (1) if $f\epsilon\mathfrak{M}$, then one and only one of the three relations $f\epsilon\mathfrak{P}$, $f\epsilon\mathfrak{N}$, $f\epsilon\mathfrak{A}$ is valid; (2) if $f\epsilon\mathfrak{M}$, $g\epsilon\mathfrak{M}$, $0\leqq f\leqq g$, and $g\epsilon\mathfrak{A}$, then $f\epsilon\mathfrak{A}$; (3) if $f\epsilon\mathfrak{P}$, $g\epsilon\mathfrak{M}$ and $f\equiv g \pmod{\mathfrak{A}}$, then $g\epsilon\mathfrak{P}$; (4) if $f\epsilon\mathfrak{P}$ and $g\epsilon\mathfrak{P}$, then $f+g\epsilon\mathfrak{P}$ and $fg\epsilon\mathfrak{P}$.

To prove (1) we proceed as follows. Since $(-f+|f|)(f+|f|)=-f^2+|f|^2=0$, the fact that $\mathfrak{M}/\mathfrak{A}$ is a field implies that at least one of the relations $-f+|f|\equiv 0 \pmod{\mathfrak{A}}$ and $+f+|f|\equiv 0 \pmod{\mathfrak{A}}$ is valid. If both hold, then $2f\equiv(f+|f|)-(-f+|f|)\equiv 0 \pmod{\mathfrak{A}}$ and therefore $f\equiv 0 \pmod{\mathfrak{A}}$. We see therefore that one and only one of the three sets of relations $f\equiv|f|\not\equiv 0 \pmod{\mathfrak{A}}$, $-f\equiv|-f|\not\equiv 0 \pmod{\mathfrak{A}}$, $f\equiv 0 \pmod{\mathfrak{A}}$ is valid, as we wished to show.

We establish (2) by contradiction. Let us suppose that $f\not\equiv 0 \pmod{\mathfrak{A}}$. Then there exists an element h in $\mathfrak{M}$ such that $fh\equiv 1 \pmod{\mathfrak{A}}$. By (1) at least one of the relations $-h\equiv|h| \pmod{\mathfrak{A}}$ and $h\equiv|h| \pmod{\mathfrak{A}}$ is valid. If the first should hold, we would have $1+f|h|\equiv 1-fh\equiv 0 \pmod{\mathfrak{A}}$, and also $1+f|h|\geqq 1$ on account of the inequality $f\geqq 0$. Here we have a contradiction: it is obvious that $1/(1+f|h|)$ is in $\mathfrak{M}$; and it follows that $1=(1+f|h|)[1/(1+f|h|)]\equiv 0 \pmod{\mathfrak{A}}$, $1\epsilon\mathfrak{A}$. If the second should hold, we would have $1-f|h|\equiv 1-fh\equiv 0 \pmod{\mathfrak{A}}$, $g|h|+(1-f|h|)\equiv 0 \pmod{\mathfrak{A}}$, and also $g|h|+(1-f|h|)\geqq 1$ on account of the inequality $f\leqq g$. Again we have a contradiction. Thus our initial assumption is false, as we wished to prove.

---

* See van der Waerden, *Moderne Algebra*, I, Leipzig, 1930, Chapter X.

We now consider (3). Since $f\equiv g \pmod{\mathfrak{A}}$ and $f\not\equiv 0 \pmod{\mathfrak{A}}$, it is obvious that $g\not\equiv 0 \pmod{\mathfrak{A}}$. If the relation $-g\equiv |g| \pmod{\mathfrak{A}}$ held, we could combine it with $f\equiv |f| \pmod{\mathfrak{A}}$ to obtain the relations $|f|+|g|\equiv f-g\equiv 0 \pmod{\mathfrak{A}}$. We could then conclude by virtue of (2) that $|g|\equiv 0 \pmod{\mathfrak{A}}$ and hence that $g\equiv g+|g|\equiv 0 \pmod{\mathfrak{A}}$. This contradiction shows that we must have $g\equiv |g|\not\equiv 0 \pmod{\mathfrak{A}}$, or $g\epsilon\mathfrak{P}$, in accordance with (1).

The proof of (4) is similar to that of (3). The relations $f\equiv |f|\not\equiv 0 \pmod{\mathfrak{A}}$ and $g\equiv |g|\not\equiv 0 \pmod{\mathfrak{A}}$ imply $fg\equiv |fg|\not\equiv 0 \pmod{\mathfrak{A}}$; in other words, $f\epsilon\mathfrak{P}$ and $g\epsilon\mathfrak{P}$ imply $fg\epsilon\mathfrak{P}$. These relations also imply that $f+g\not\equiv -|f+g| \pmod{\mathfrak{A}}$: for otherwise we would have $|f|+|g|+|f+g|\equiv f+g-(f+g)\equiv 0 \pmod{\mathfrak{A}}$ and hence, by virtue of (2), $|f|\equiv |g|\equiv 0 \pmod{\mathfrak{A}}$, contrary to hypothesis. Furthermore, these relations show that $f+g\not\equiv 0 \pmod{\mathfrak{A}}$: for otherwise we would have $|f|+|g|\equiv f+g\equiv 0 \pmod{\mathfrak{A}}$ and hence, by virtue of (2), $|f|\equiv |g|\equiv 0 \pmod{\mathfrak{A}}$, contrary to hypothesis. It now follows that $f+g\equiv |f+g|\not\equiv 0 \pmod{\mathfrak{A}}$ or, equivalently, that $f+g\epsilon\mathfrak{P}$, as we wished to prove.

If we remark that the relations $\alpha\equiv |\alpha|\not\equiv 0 \pmod{\mathfrak{A}}$ and $\alpha=|\alpha|\neq 0$ are equivalent whenever $\alpha\epsilon R$, we see that the partition of $\mathfrak{M}$ into the disjoint classes $\mathfrak{P}$, $\mathfrak{N}$, and $\mathfrak{A}$ has all the properties required above. It defines a corresponding partition of the field $\mathfrak{M}/\mathfrak{A}$; and this partition of $\mathfrak{M}/\mathfrak{A}$ in turn defines a simple ordering of the field, in a familiar way.* We can now prove that the order so introduced is an archimedean order; in other words, that $f\epsilon\mathfrak{P}$ implies that $nf-1\epsilon\mathfrak{P}$ for some integer n. If the element $f-1/n$ should belong to $\mathfrak{P}$ for no integer n, we would have $-(f-1/n)\equiv |f-1/n| \pmod{\mathfrak{A}}$ or, equivalently, $(f-1/n)+|f-1/n|\epsilon\mathfrak{A}$ for every n. On passing to the limit, we would then obtain $f+|f|\epsilon\mathfrak{A}^-=\mathfrak{A}$ or, equivalently, $f\bar{\epsilon}\mathfrak{P}$. This contradiction shows that $f-1/n\epsilon\mathfrak{P}$ for some integer n. By (4), we conclude that $nf-1=n(f-1/n)\epsilon\mathfrak{P}$ for that integer, as we wished. Since the field $\mathfrak{M}/\mathfrak{A}$ has an archimedean ordering which is an extension of the natural ordering of its subfield R, it must coincide with R.

Finally we consider the special properties of the homomorphism $\mathfrak{M}\rightarrow R$ stated in the theorem. In order that this homomorphism take f into α it is obviously necessary and sufficient that $f\equiv\alpha \pmod{\mathfrak{A}}$. If $f\equiv\alpha \pmod{\mathfrak{A}}$, the relation $|f-\alpha|\geqq\epsilon>0$ is impossible: for $f\equiv\alpha \pmod{\mathfrak{A}}$ implies $f-\alpha\equiv 0 \pmod{\mathfrak{A}}$, $|f-\alpha|\equiv 0 \pmod{\mathfrak{A}}$; and $|f-\alpha|\geqq\epsilon>0$ implies $|f-\alpha|\not\equiv 0 \pmod{\mathfrak{A}}$ in accordance with (2) above. It follows that g.l.b. $f\leqq\alpha\leqq$ l.u.b. f; or, more precisely, that α is a limit point of the range of f. The proof that $f\equiv\alpha \pmod{\mathfrak{A}}$ implies $|f|\equiv|\alpha| \pmod{\mathfrak{A}}$ is simple and will be omitted.

We pass now to the investigation of the connections between CR-spaces and T_0-spaces. The fundamental theorem reads as follows.

* See van der Waerden, *Moderne Algebra*, I, Leipzig, 1930, p. 209.

THEOREM 77. *In a T_0-space $\mathfrak{R}$, let $\mathcal{G}$ be the family of all open sets $\mathfrak{G}=\mathfrak{R}(\alpha<f<\beta)$ where $f\epsilon\mathfrak{M}$ and $\alpha\epsilon R$, $\beta\epsilon R$; let $\mathfrak{X}(\mathfrak{r})$ be the intersection of all those sets in $\mathcal{G}$ which contain the point $\mathfrak{r}$ in $\mathfrak{R}$; let $\mathfrak{X}$ be the family of all sets $\mathfrak{X}(\mathfrak{r})$; and let R^* be the space obtained by assigning each subset of $\mathfrak{X}$ specified by the relations $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{G}$, $\mathfrak{G}\epsilon\mathcal{G}$ as a neighborhood of every $\mathfrak{X}(\mathfrak{r})$ which it contains. Then $\mathfrak{R}^*$ is a CR-space which is a continuous image of $\mathfrak{R}$ under the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$; and the function-rings $\mathfrak{M}$ and $\mathfrak{M}^*$ for $\mathfrak{R}$ and $\mathfrak{R}^*$ respectively are analytically isomorphic under the correspondence $f\rightarrow f^*$ defined by setting $f^*(\mathfrak{X}(\mathfrak{r}))=f(\mathfrak{r})$ for each point $\mathfrak{r}$ in $\mathfrak{R}$. In case $\mathfrak{R}$ is itself a CR-space, $\mathfrak{R}^*$ is topologically equivalent to $\mathfrak{R}$.*

If $\mathfrak{r}$ is an arbitrary point and f an arbitrary function in $\mathfrak{M}$, the set $\mathfrak{R}(\alpha<f<\beta)$ contains $\mathfrak{r}$ if α and β are so chosen that $\alpha<f(\mathfrak{r})<\beta$. Hence every point $\mathfrak{r}$ determines a set $\mathfrak{X}(\mathfrak{r})$. It is easily seen that every function in $\mathfrak{M}$ is constant on each set $\mathfrak{X}(\mathfrak{r})$. In fact, if $f(\mathfrak{s})\neq f(\mathfrak{r})$, α and β could be chosen so that $\mathfrak{R}(\alpha<f<\beta)$ contains $\mathfrak{r}$ and $\mathfrak{X}(\mathfrak{r})$ but not $\mathfrak{s}$. It is therefore clear that $\mathfrak{X}(\mathfrak{r})\mathfrak{X}(\mathfrak{s})\neq 0$ implies $\mathfrak{X}(\mathfrak{r})=\mathfrak{X}(\mathfrak{s})$: for every function in $\mathfrak{M}$ must be constant on $\mathfrak{X}(\mathfrak{r})\cup\mathfrak{X}(\mathfrak{s})$; and every set $\mathfrak{R}(\alpha<f<\beta)$, $f\epsilon\mathfrak{M}$, must contain both or neither of the sets $\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{X}(\mathfrak{s})$. From these results it is obvious that every set in $\mathcal{G}$ is the union of all the sets $\mathfrak{X}(\mathfrak{r})$ which it contains. Hence the assignment of neighborhoods described in the statement of the theorem can be justified through the properties which follow: (1) every $\mathfrak{X}(\mathfrak{r})$ has at least one neighborhood; (2) the intersection of two neighborhoods of $\mathfrak{X}(\mathfrak{r})$ is a neighborhood of $\mathfrak{X}(\mathfrak{r})$; (3) any neighborhood which contains $\mathfrak{X}(\mathfrak{r})$ is a neighborhood of $\mathfrak{X}(\mathfrak{r})$; (4) if $\mathfrak{X}(\mathfrak{r})\neq\mathfrak{X}(\mathfrak{s})$, there is a neighborhood of $\mathfrak{X}(\mathfrak{r})$ which does not contain $\mathfrak{X}(\mathfrak{s})$. The properties (1), (3), (4) are obvious from preceding remarks, while (2) follows at once from the observation† that $\mathfrak{R}(\alpha<f<\beta)\mathfrak{R}(\gamma<g<\delta)=\mathfrak{R}(0<h<\lambda)$ where $h=\min\ [(f-\alpha)(\beta-f),\ (g-\gamma)(\delta-g)]$, $\lambda>\|h\|$. Accordingly the introduction of the indicated neighborhood-system yields a T_1-space $\mathfrak{R}^*$. Since the sets $\mathfrak{X}(\mathfrak{r})$ are disjoint, the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$ from $\mathfrak{R}$ to $\mathfrak{R}^*$ is univocal. The open sets specified in $\mathfrak{R}^*$ by the relations $\mathfrak{X}(\mathfrak{r})\subset\mathfrak{G}$, $\mathfrak{G}\epsilon\mathcal{G}$ constitute a basis for $\mathfrak{R}^*$ and have as antecedents in $\mathfrak{R}$ the corresponding sets $\mathfrak{G}$ in $\mathcal{G}$. The correspondence $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$ is therefore continuous. Furthermore, if $\mathfrak{R}$ is a CR-space, the family $\mathcal{G}$ is a basis for $\mathfrak{R}$, so that $\mathfrak{X}(\mathfrak{r})=\{\mathfrak{r}\}$ and the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$ is biunivocal and bicontinuous. Thus $\mathfrak{R}^*$ is a continuous image of $\mathfrak{R}$, and is topologically equivalent to $\mathfrak{R}$ when $\mathfrak{R}$ is a CR-space. If $f\epsilon\mathfrak{M}$, the function f^* defined in $\mathfrak{R}^*$ by putting $f^*(\mathfrak{X}(\mathfrak{r}))=f(\mathfrak{r})$ is bounded, single-valued, and real. We see also that f^* is continuous since the set

† We use the relations $2\max(f, g)=|f-g|+f+g$, $2\min(f, g)=-|f-g|+f+g$ here and elsewhere in this section, when properties of $\max(f, g)$ or $\min(f, g)$ are needed.

$\mathfrak{R}^*(\alpha < f^* < \beta)$ is the open set of all $\mathfrak{X}(\mathfrak{r})$ such that $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{R}(\alpha < f < \beta)$. Thus f^* is in $\mathfrak{M}^*$. On the other hand if f^* is any function in $\mathfrak{M}^*$, the function $f(\mathfrak{r}) = f^*(\mathfrak{X}(\mathfrak{r}))$ is a bounded real function; and $f(\mathfrak{r})$ is also continuous since it defines a correspondence $\mathfrak{r} \to f(\mathfrak{r})$ from $\mathfrak{R}$ to R which is obtained by eliminating $\mathfrak{X}(\mathfrak{r})$ from the continuous correspondences $\mathfrak{r} \to \mathfrak{X}(\mathfrak{r})$, $\mathfrak{X}(\mathfrak{r}) \to f^*(\mathfrak{X}(\mathfrak{r}))$ carrying $\mathfrak{R}$ into $\mathfrak{R}^*$ and $\mathfrak{R}^*$ into R respectively. Thus the correspondence $f \to f^*$ takes $\mathfrak{M}$ univocally into $\mathfrak{M}^*$. It is easily verified that this correspondence is an analytical isomorphism in accordance with Definition 22. Finally, it is evident that $\mathfrak{R}^*$ is a CR-space, since the sets $\mathfrak{R}^*(\alpha < f^* < \beta)$ constitute a basis for $\mathfrak{R}^*$ by virtue of the characterization given for them above. We may note in passing that $\mathfrak{X}(\mathfrak{r})$ can obviously be obtained as the intersection of the closed sets $\mathfrak{R}(f = f(\mathfrak{r}))$, $f \epsilon \mathfrak{M}$, and is therefore closed.

The immediate significance of Theorem 77 is seen to be that in studying function-rings we may restrict attention to the case of CR-spaces without any loss of generality.

We shall now obtain some information concerning the connections between function-rings and algebraic maps for CR-spaces.

THEOREM 78. *In a CR-space* $\mathfrak{R}$, *let* $\mathcal{G}$ *be the basis of all sets* $\mathfrak{R}(\alpha < f < \beta)$ *where f is in the function-ring* $\mathfrak{M}$; *let A be the basic ring generated by* $\mathcal{G}$; *let* $\mathfrak{A}$ *be a divisorless ideal in* $\mathfrak{M}$; *and let* $\mathfrak{a}(\mathfrak{A})$ *be the ideal in A determined as the class of all sets a in A such that, for at least one choice of f,* α, *and* β, *the relations* $a < \mathfrak{R}(\alpha < f < \beta)$, $f \equiv \gamma \pmod{\mathfrak{A}}$, *and either* $\gamma < \alpha$ *or* $\gamma > \beta$ *are valid. Then in the bicompact Boolean space* $\mathfrak{E}(A)$, *the closed sets* $\mathfrak{Z}(\mathfrak{A}) = \mathfrak{E}'(\mathfrak{a}(\mathfrak{A}))$ *are disjoint and constitute a continuous covering family* Z. *Under the usual topology,* Z *defines a bicompact H-space* $\mathfrak{Q}$.

We must first show that $\mathfrak{a}(\mathfrak{A})$ is an ideal in A. It is evident that $0 \epsilon \mathfrak{a}(\mathfrak{A})$ and that $a \epsilon \mathfrak{a}(\mathfrak{A})$ and $a > c$ imply $c \epsilon \mathfrak{a}(\mathfrak{A})$. Hence we have only to prove that $a \epsilon \mathfrak{a}(\mathfrak{A})$ and $b \epsilon \mathfrak{a}(\mathfrak{A})$ imply $a \vee b \epsilon \mathfrak{a}(\mathfrak{A})$. By hypothesis, we have $a < \mathfrak{R}(\alpha < f < \beta)$ and $b < \mathfrak{R}(\gamma < g < \delta)$ where $f \equiv \sigma \pmod{\mathfrak{A}}$, $g \equiv \tau \pmod{\mathfrak{A}}$, and σ and τ lie respectively outside the closed intervals $[\alpha, \beta]$ and $[\gamma, \delta]$. If we put $h = |f - \sigma| + |g - \tau|$, $\epsilon > 0$, $\eta > \|h\|$, and choose ϵ sufficiently small, it is seen that $a \vee b < \mathfrak{R}(h > \epsilon) = \mathfrak{R}(\epsilon < h < \eta)$, $h \equiv 0 \pmod{\mathfrak{A}}$ and hence that $a \vee b \epsilon \mathfrak{a}(\mathfrak{A})$, as we wished to prove.

Next we show that $\mathfrak{A}_1 \neq \mathfrak{A}_2$ implies $\mathfrak{a}(\mathfrak{A}_1) \vee \mathfrak{a}(\mathfrak{A}_2) = \mathfrak{e}$ and hence $\mathfrak{Z}(\mathfrak{A}_1)\mathfrak{Z}(\mathfrak{A}_2) = \mathfrak{E}'(\mathfrak{a}(\mathfrak{A}_1))\mathfrak{E}'(\mathfrak{a}(\mathfrak{A}_2)) = \mathfrak{E}'(\mathfrak{a}(\mathfrak{A}_1) \vee \mathfrak{a}(\mathfrak{A}_2)) = \mathfrak{E}'(\mathfrak{e}) = 0$. If $\mathfrak{A}_1 \neq \mathfrak{A}_2$, the ideal generated in $\mathfrak{M}$ by $\mathfrak{A}_1$ and $\mathfrak{A}_2$ coincides with $\mathfrak{M}$; in other words, there exist functions f and g such that $f + g = 1$, $f \epsilon \mathfrak{A}_1$, $g \epsilon \mathfrak{A}_2$. The relations $|f| \equiv 0 \pmod{\mathfrak{A}_1}$, $|g| \equiv 0 \pmod{\mathfrak{A}_2}$ show that the sets $a = \mathfrak{R}(|f| > 1/3)$ and $b = \mathfrak{R}(|g| > 1/3)$ belong to $\mathfrak{a}(\mathfrak{A}_1)$ and to $\mathfrak{a}(\mathfrak{A}_2)$ respectively. Since $1 = f + g \leqq |f| + |g|$, we must

have $a'b' = \mathfrak{R}(|f| \leqq 1/3) \cdot \mathfrak{R}(|g| \leqq 1/3) = 0$ or, equivalently, $a \vee b = e$. It follows that $\mathfrak{a}(\mathfrak{A}_1) \vee \mathfrak{a}(\mathfrak{A}_2) = \mathfrak{e}$, as we wished to prove.

We must also show that, if $\mathfrak{p}$ is any prime ideal in A, there exists a divisorless ideal $\mathfrak{A}$ in $\mathfrak{M}$ for which the equivalent relations $\mathfrak{p} \epsilon \mathfrak{Z}(\mathfrak{A})$, $\mathfrak{p} \epsilon \mathfrak{E}(\mathfrak{a}(\mathfrak{A}))$, $\mathfrak{a}(\mathfrak{A}) \subset \mathfrak{p}$ are valid. By Theorem 76, the construction of a divisorless ideal is equivalent to the construction of a homomorphism $\mathfrak{M} \rightarrow R$. Now, if f is a fixed function in $\mathfrak{M}$, we may classify the open intervals (α, β), $-\infty < \alpha < \beta < +\infty$, according to the relation of the set $\mathfrak{R}(\alpha < f < \beta)$ to the given prime ideal $\mathfrak{p}$ in A; in particular, we consider the intervals (α, β) for which $\mathfrak{R}(\alpha < f < \beta) \epsilon \mathfrak{p}$. We prove that there is exactly one real number which belongs to none of the latter intervals. First let us show that there is at least one. If there were not, the open intervals (α, β) under consideration would cover the closed interval $[\alpha', \beta']$ where α' and β' are bounds for f. By the Heine-Borel-Lebesgue covering property, a finite number of the intervals (α, β) would suffice to cover $[\alpha', \beta']$; and in consequence a finite number of the sets $\mathfrak{R}(\alpha < f < \beta)$ in $\mathfrak{p}$ would suffice to cover $\mathfrak{R}$. Here we have a contradiction, since no finite union of sets belonging to $\mathfrak{p}$ can coincide with the unit $\mathfrak{R}$ in A. Thus there exists at least one real number γ of the desired kind; and it is evident that any such number γ must satisfy the inequality g.l.b. $f \leqq \gamma \leqq$ l.u.b. f or, more precisely, must belong to the closure of the range of f. Now let us prove that, if γ_1 and γ_2 are such numbers, then $\gamma_1 = \gamma_2$. By hypothesis, the sets $\mathfrak{R}(\gamma_1 - \epsilon < f < \gamma_1 + \epsilon)$ and $\mathfrak{R}(\gamma_2 - \epsilon < f < \gamma_2 + \epsilon)$ belong to $\mathfrak{p}$ for no positive ϵ. It follows that these sets must have non-void intersection for every positive ϵ and hence that the open intervals $(\gamma_1 - \epsilon, \gamma_1 + \epsilon)$ and $(\gamma_2 - \epsilon, \gamma_2 + \epsilon)$ must have points in common for every positive ϵ. Thus we conclude that $\gamma_1 = \gamma_2$. In view of the preceding construction, we see that the prime ideal $\mathfrak{p}$ assigns to each f in $\mathfrak{M}$ a unique corresponding real number γ. We wish now to prove that the correspondence $f \rightarrow \gamma$ defines a homomorphism $\mathfrak{M} \rightarrow R$. By construction, it is evident that $f \rightarrow \gamma$ implies $f - \gamma \rightarrow 0$. It is also evident that, if f is constant, then $f \rightarrow \gamma$ if and only if $f = \gamma$. If we can establish the propositions (1) $f \rightarrow 0$, $g \rightarrow 0$ imply $f + g \rightarrow 0$, and (2) $f \rightarrow 0$ implies $fg \rightarrow 0$, we can then show that $f \rightarrow \gamma$, $g \rightarrow \delta$ imply $(f - \gamma) + (g - \delta) \rightarrow 0$ or, equivalently, $f + g \rightarrow \gamma + \delta$ and $(f - \gamma)g + \gamma(g - \delta) \rightarrow 0$ or, equivalently, $fg \rightarrow \gamma\delta$. If $f \rightarrow 0$ and $g \rightarrow 0$, the sets $\mathfrak{R}(|f| < \epsilon/2)$ and $\mathfrak{R}(|g| < \epsilon/2)$ belong to $\mathfrak{p}$ for no positive ϵ; and their intersection likewise belongs to $\mathfrak{p}$ for no positive ϵ. Since $\mathfrak{R}(|f+g| < \epsilon) \supset \mathfrak{R}(|f| < \epsilon/2)\mathfrak{R}(|g| < \epsilon/2)$, the set $\mathfrak{R}(|f+g| < \epsilon)$ belongs to $\mathfrak{p}$ for no positive ϵ. It follows that $f + g \rightarrow 0$. Similarly if $f \rightarrow 0$, the set $\mathfrak{R}(|fg| < \epsilon)$ contains the set $\mathfrak{R}(|f| < \epsilon/\|g\|)$ which does not belong to $\mathfrak{p}$ for any positive ϵ; and we conclude that $fg \rightarrow 0$. Thus the correspondence $f \rightarrow \gamma$ is a homomorphism.

If $\mathfrak{A}$ is the ideal determined by this homomorphism, we must show that

$\mathfrak{a}(\mathfrak{A}) \subset \mathfrak{p}$. If $a \epsilon \mathfrak{a}(\mathfrak{A})$, there exists a function f and real numbers α and β such that $a < \mathfrak{R}(\alpha < f < \beta)$ and $f \to \gamma$, where γ is outside the closed interval $[\alpha, \beta]$. For sufficiently small positive ϵ we have $\mathfrak{R}(\gamma - \epsilon < f < \gamma + \epsilon)\mathfrak{R}(\alpha < f < \beta) = 0$. Since, by our construction of γ, the set $\mathfrak{R}(\gamma - \epsilon < f < \gamma + \epsilon)$ does not belong to $\mathfrak{p}$, we must therefore have $\mathfrak{R}(\alpha < f < \beta) \epsilon \mathfrak{p}$. It follows that $a \epsilon \mathfrak{p}$ and that $\mathfrak{a}(\mathfrak{A}) \subset \mathfrak{p}$.

The family Z is now seen to cover $\mathfrak{C}(A)$; but we wish to show further that Z is continuous. If a is any element in A such that $\mathfrak{C}(a) \supset \mathfrak{Z}(\mathfrak{A})$, we have to find an element b in A such that $\mathfrak{Z}(\mathfrak{A}) \subset \mathfrak{C}(b)$ and such that $\mathfrak{Z}(\mathfrak{B})\mathfrak{C}'(a) \neq 0$ implies $\mathfrak{Z}(\mathfrak{B})\mathfrak{C}(b) = 0$. The relations $\mathfrak{C}(a) \supset \mathfrak{Z}(\mathfrak{A})$, $\mathfrak{C}(a') \subset \mathfrak{C}(\mathfrak{a}(\mathfrak{A}))$, and $a' \epsilon \mathfrak{a}(\mathfrak{A})$ are equivalent. Hence we can find f, α, and β so that $a' < \mathfrak{R}(\alpha < f < \beta)$ where $f \to \gamma$ and γ is outside the closed interval $[\alpha, \beta]$; and we then see that $a > \mathfrak{R}(|g| \leqq \epsilon)$ where $g = f - \gamma$ and ϵ is a sufficiently small positive number. We now choose b as the set $\mathfrak{R}(|g| \leqq \epsilon/3)$. Since $b' = \mathfrak{R}(|g| > \epsilon/3)$ where $|g| \equiv 0 \pmod{\mathfrak{A}}$, we see that $b' \epsilon \mathfrak{a}(\mathfrak{A})$ and hence that $\mathfrak{C}(b) \supset \mathfrak{Z}(\mathfrak{A})$. If $\mathfrak{B}$ is a divisorless ideal in $\mathfrak{M}$ such that $\mathfrak{Z}(\mathfrak{B})\mathfrak{C}(a') = \mathfrak{Z}(\mathfrak{B})\mathfrak{C}'(a) \neq 0$ and if the homomorphism $\mathfrak{M} \to \mathfrak{M}/\mathfrak{B}$ carries g into δ, we introduce the set $c = \mathfrak{R}(|g - \delta| \leqq \epsilon/3)$ in A. Since $c' = \mathfrak{R}(|g - \delta| > \epsilon/3)$ where $|g - \delta| \equiv 0 \pmod{\mathfrak{B}}$, we have $c' \epsilon \mathfrak{a}(\mathfrak{B})$, $\mathfrak{C}(c) \supset \mathfrak{Z}(\mathfrak{B})$. The relations $\mathfrak{C}(a'c) = \mathfrak{C}'(a)\mathfrak{C}(c) \supset \mathfrak{C}'(a)\mathfrak{Z}(\mathfrak{B}) \neq 0$ show that $a'c \neq 0$. From the relations $\mathfrak{R}(|g| > \epsilon)\mathfrak{R}(|g - \delta| \leqq \epsilon/3) \supset a'c \neq 0$, we infer that $|\delta| > 2\epsilon/3$. We then conclude that $bc = \mathfrak{R}(|g| \leqq \epsilon/3)\mathfrak{R}(|g - \delta| \leqq \epsilon/3) = 0$ and hence that $\mathfrak{C}(b)\mathfrak{Z}(\mathfrak{B}) \subset \mathfrak{C}(b)\mathfrak{C}(c) = \mathfrak{C}(bc) = \mathfrak{C}(0) = 0$, $\mathfrak{C}(b)\mathfrak{Z}(\mathfrak{B}) = 0$, as we wished to prove.

If we now impose the usual topology upon the family Z, we obtain an H-space $\mathfrak{Q}$ and a map $m(\mathfrak{Q}, \mathfrak{C}(A), \mathrm{Z})$ in accordance with Theorem 23. Since Z is a continuous covering family, Theorem 22 shows that $\mathfrak{Q}$ is a continuous image of $\mathfrak{C}(A)$. It follows that $\mathfrak{Q}$ is a bicompact H-space.

THEOREM 79. *The space $\mathfrak{Q}$ of Theorem 78 is an immediate, and hence strict, H-extension of the given CR-space $\mathfrak{R}$. Every function in $\mathfrak{M}$ can be extended from $\mathfrak{R}$ to $\mathfrak{Q}$ so as to be continuous, and hence bounded, in $\mathfrak{Q}$. If $f \epsilon \mathfrak{M}$ and f^* is its extension to $\mathfrak{Q}$, then the correspondence $f \to f^*$ is an analytical isomorphism between the function-rings for $\mathfrak{R}$ and $\mathfrak{Q}$. In case $\mathfrak{R}$ is a bicompact H-space, $\mathfrak{Q}$ coincides with $\mathfrak{R}$.*

Considering the algebraic map $m(\mathfrak{R}, \mathfrak{C}(A), \mathfrak{X})$, where A is the basic ring described in Theorem 78, we show that each set $\mathfrak{X}(\mathfrak{r})$ coincides with a suitable member of the family Z of that theorem. If $\mathfrak{r}$ is an arbitrary point in $\mathfrak{R}$, the correspondence $f \to f(\mathfrak{r})$ defines a homomorphism $\mathfrak{M} \to R$. The associated divisorless ideal $\mathfrak{A} = \mathfrak{A}(\mathfrak{r})$ then has the property that $\mathfrak{X}(\mathfrak{r}) = \mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$. We prove this statement as follows. If $a \epsilon A$ and $\mathfrak{C}(a) \supset \mathfrak{X}(\mathfrak{r})$, then the basis $\mathcal{G}$ contains a set $b = \mathfrak{R}(|f| < \epsilon)$, $f(\mathfrak{r}) = 0$, which contains $\mathfrak{r}$ and is contained in a: for Theorem 28 shows that $\mathfrak{r}$ is interior to a. The relations $b' = \mathfrak{R}(|f| \geqq \epsilon) \subset \mathfrak{R}(|f| > \epsilon/2)$ and

$f\equiv 0$ (mod $\mathfrak{A}(\mathfrak{r})$) show that $b'\epsilon\mathfrak{a}(\mathfrak{A}(\mathfrak{r}))$ or, equivalently, that $\mathfrak{E}(b)\supset\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$. Since $\mathfrak{E}(a)\supset\mathfrak{E}(b)$, we conclude that $\mathfrak{E}(a)$ also contains $\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$. On the other hand, if $a\epsilon A$ and $\mathfrak{E}(a)\supset\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$, we know that $a'\epsilon\mathfrak{a}(\mathfrak{A}(\mathfrak{r}))$. Hence there exist a function f and a positive number ϵ such that $a'<\mathfrak{R}(|f|>\epsilon)$ where $f\equiv 0$ (mod $\mathfrak{A}(\mathfrak{r})$) or, equivalently, $f(\mathfrak{r})=0$. Consequently the open set $b=\mathfrak{R}(|f|<\epsilon)$ belongs to A, contains $\mathfrak{r}$, and is contained in a. Since $\mathfrak{r}$ is thus interior to a, we conclude that $\mathfrak{E}(a)\supset\mathfrak{X}(\mathfrak{r})$. The equivalence of the relations $\mathfrak{E}(a)\supset\mathfrak{X}(\mathfrak{r})$, $\mathfrak{E}(a)\supset\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$ implies that $\mathfrak{X}(\mathfrak{r})=\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$, as we wished to show. Theorem 41 now shows that $\mathfrak{Q}$ is an immediate H-extension of $\mathfrak{R}$; and Theorems 64 and 75 show that $\mathfrak{Q}$ must be a strict extension of $\mathfrak{R}$.

In discussing the extension of functions from $\mathfrak{R}$ to $\mathfrak{Q}$ we may regard $\mathfrak{R}$ as a subspace of $\mathfrak{Q}$ and we may even identify $\mathfrak{R}$ and $\mathfrak{Q}$ with the topologized families X and Z respectively. If $f\epsilon\mathfrak{M}$, we define the extended function f^* by putting $f^*(\mathfrak{Z}(\mathfrak{A}))=\gamma$ where $f\rightarrow\gamma$ under the homomorphism $\mathfrak{M}\rightarrow\mathfrak{M}/\mathfrak{A}$. That f^* is actually an extension of f appears at once from the relations $f^*(\mathfrak{X}(\mathfrak{r}))=f^*(\mathfrak{Z}(\mathfrak{A}(\mathfrak{r})))=f(\mathfrak{r})$. We see also that $f^*=g^*$ if and only if $f=g$, and that $(f+g)^*=f^*+g^*$, $(fg)^*=f^*g^*$, $|f|^*=|f^*|$. It is evident that $\|f\|\leqq\|f^*\|$; and we know that $f\rightarrow\gamma$ implies $|\gamma|\leqq\|f\|$. Hence we see that $\|f\|=\|f^*\|$. To show that f^* is continuous, we have to prove that $\mathfrak{Q}(\alpha<f^*<\beta)$ is an open set. If $\mathfrak{Z}(\mathfrak{A}_0)$ represents a point $\mathfrak{q}_0$ in this set, then $f^*(\mathfrak{Z}(\mathfrak{A}_0))=\gamma_0$ where $f\rightarrow\gamma_0$ under the homomorphism $\mathfrak{M}\rightarrow\mathfrak{M}/\mathfrak{A}_0$ and γ_0 satisfies the inequality $\alpha<\gamma_0<\beta$. Putting $0<\epsilon<\min(\gamma_0-\alpha,\beta-\gamma_0)$ we consider the set $a=\mathfrak{R}(\gamma_0-\epsilon<f<\gamma_0+\epsilon)$ $=\mathfrak{R}(|f-\gamma_0|<\epsilon)$ in A. We shall show that the open set specified in $\mathfrak{Q}$ by the relation $\mathfrak{Z}(\mathfrak{A})\subset\mathfrak{E}(a)$ contains $\mathfrak{q}_0$ and is contained in $\mathfrak{Q}(\alpha<f^*<\beta)$; and we can then infer that $\mathfrak{Q}(\alpha<f^*<\beta)$ is open. First we prove that $\mathfrak{Z}(\mathfrak{A}_0)\subset\mathfrak{E}(a)$. The set $b=\mathfrak{R}(|f-\gamma_0|>\epsilon/2)$ is in A and obviously contains $a'=\mathfrak{R}(|f-\gamma_0|\geqq\epsilon)$. Since $f-\gamma_0\equiv 0$ (mod $\mathfrak{A}_0$), we see that $b\epsilon\mathfrak{a}(\mathfrak{A}_0)$ and hence that $a'\epsilon\mathfrak{a}(\mathfrak{A}_0)$. It follows that $\mathfrak{E}(a)\supset\mathfrak{Z}(\mathfrak{A}_0)$, as we wished to show. Secondly, we prove that $\mathfrak{Z}(\mathfrak{A})\subset\mathfrak{E}(a)$ implies $f^*(\mathfrak{Z}(\mathfrak{A}))=\gamma$ where $\alpha<\gamma<\beta$; and we can then infer that the point $\mathfrak{q}$ represented by $\mathfrak{Z}(\mathfrak{A})$ is in $\mathfrak{Q}(\alpha<f^*<\beta)$. Now $\mathfrak{Z}(\mathfrak{A})\subset\mathfrak{E}(a)$ implies that $\mathfrak{R}(|f-\gamma_0|\geqq\epsilon)=a'\epsilon\mathfrak{a}(\mathfrak{A})$. Hence, if $f\rightarrow\gamma$ under the homomorphism $\mathfrak{M}\rightarrow\mathfrak{M}/\mathfrak{A}$, we must have $|\gamma-\gamma_0|\leqq\epsilon$: for the relations $a\vee a'=e$, $a'\epsilon\mathfrak{a}(\mathfrak{A})$, and $\mathfrak{a}(\mathfrak{A})\neq\mathfrak{e}$ show that a is not in $\mathfrak{a}(\mathfrak{A})$ and hence that $\gamma-\gamma_0$ is not outside the closed interval $[-\epsilon,\epsilon]$. Since $f\rightarrow\gamma$ implies $f^*(\mathfrak{Z}(\mathfrak{A}))=\gamma$ and since $\alpha<\gamma_0-\epsilon\leqq\gamma$ $\leqq\gamma_0+\epsilon<\beta$, our proof is brought to the desired conclusion.

If an arbitrary bounded continuous real function f^* in $\mathfrak{Q}$ is restricted to the subspace $\mathfrak{R}$, the restricted or partial function f obviously belongs to $\mathfrak{M}$. Since $\mathfrak{R}$ is everywhere dense in $\mathfrak{Q}$, the extension of f to $\mathfrak{Q}$ must be the original function f^*. We conclude therefore that the correspondence $f\rightarrow f^*$ defines an analytical isomorphism of the function-rings $\mathfrak{M}$ and $\mathfrak{M}^*$ for $\mathfrak{R}$ and $\mathfrak{Q}$ respectively.

If the space $\mathfrak{R}$ is a bicompact H-space, it is normal and hence completely regular. The corresponding space $\mathfrak{Q}$ must then coincide with $\mathfrak{R}$, since a bicompact H-space is absolutely closed with respect to immediate or strict H-extension.

Obviously Theorem 79 provides a further reduction of the theory of function-rings: it enables us to restrict attention to the case where the underlying T_0-space $\mathfrak{R}$ is a bicompact H-space. We shall continue the analysis of function-rings under this assumption. First we state a result which is a direct corollary of Theorem 79.

THEOREM 80. *If $\mathfrak{A}$ is a divisorless ideal in the function-ring for a bicompact H-space $\mathfrak{R}$, then there exists a point $\mathfrak{r}$ in $\mathfrak{R}$ such that $f\epsilon\mathfrak{A}$ if and only if $f(\mathfrak{r})=0$; the homomorphism $\mathfrak{M}\rightarrow\mathfrak{M}/\mathfrak{A}$ carries f into $f(\mathfrak{r})$.*

The proof of Theorem 79 shows that the families X and Z coincide when $\mathfrak{R}$ is a bicompact H-space; and the result stated here then follows from the relation $\mathfrak{X}(\mathfrak{r})=\mathfrak{Z}(\mathfrak{A}(\mathfrak{r}))$ previously noted. A direct proof can also be given. Since $\mathfrak{R}$ is bicompact, the range of any function f in $\mathfrak{M}$ is a bounded closed set of real numbers. Hence $f\epsilon\mathfrak{A}$ implies that $f\rightarrow 0$ under the homomorphism $\mathfrak{M}\rightarrow\mathfrak{M}/\mathfrak{A}$ and that the set $\mathfrak{R}(f=0)$ is non-void. If $f_1, \cdots, f_n$ are in $\mathfrak{A}$ we therefore have $\mathfrak{R}(f_1=0)\cdots\mathfrak{R}(f_n=0)=\mathfrak{R}(|f_1|=0)\cdots\mathfrak{R}(|f_n|=0)=(\mathfrak{R}|f_1|+\cdots+|f_n|=0)\neq 0$, since $|f_1|+\cdots+|f_n|\epsilon\mathfrak{A}$. We see therefore that, because of the bicompactness of $\mathfrak{R}$, the intersection of all the closed sets $\mathfrak{R}(f=0)$, where $f\epsilon\mathfrak{A}$, is non-void. It is now evident that the divisorless ideal $\mathfrak{A}(\mathfrak{r})$ defined by the homomorphic correspondence $f\rightarrow f(\mathfrak{r})$, where $\mathfrak{r}$ is a point of this intersection, contains $\mathfrak{A}$. Hence $\mathfrak{A}=\mathfrak{A}(\mathfrak{r})$, as we wished to show.

We pass now to a study of subrings of $\mathfrak{M}$.

THEOREM 81. *If $\mathfrak{M}$ is the function-ring for a bicompact H-space $\mathfrak{R}$; if $\mathfrak{N}$ is an arbitrary non-void subclass of $\mathfrak{M}$; if $\mathcal{G}_{\mathfrak{N}}$ is the family of open sets $\mathfrak{R}(\alpha<f<\beta)$, $f\epsilon\mathfrak{N}$; if $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ is the intersection of all sets $\mathfrak{G}_{\mathfrak{N}}$ in $\mathcal{G}_{\mathfrak{N}}$ which contain the point $\mathfrak{r}$ in $\mathfrak{R}$; if $\mathsf{X}_{\mathfrak{N}}$ is the family of all sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$; and if $\mathfrak{R}_{\mathfrak{N}}$ is the space obtained by assigning each subset of $\mathsf{X}_{\mathfrak{N}}$ specified by a relation $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{H}_{\mathfrak{N}}$, where $\mathfrak{H}_{\mathfrak{N}}=\mathfrak{G}_{\mathfrak{N}}^{(1)}\cdots\mathfrak{G}_{\mathfrak{N}}^{(n)}$ and $\mathfrak{G}_{\mathfrak{N}}^{(k)}\epsilon\mathcal{G}_{\mathfrak{N}}$ for $k=1,\cdots,n$, as a neighborhood of every $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ which it contains;—then $\mathfrak{R}_{\mathfrak{N}}$ is a bicompact H-space which is a continuous image of $\mathfrak{R}$. If $\mathfrak{N}_1\supset\mathfrak{N}_2$, then $\mathfrak{R}_{\mathfrak{N}_2}$ is a continuous image of $\mathfrak{R}_{\mathfrak{N}_1}$; and $\mathfrak{R}_{\mathfrak{M}}$ is topologically equivalent to $\mathfrak{R}$. If $\mathfrak{N}^*$ is the analytical subring of $\mathfrak{M}$ generated by $\mathfrak{N}$, then $\mathfrak{R}_{\mathfrak{N}^*}$ is topologically equivalent to $\mathfrak{R}_{\mathfrak{N}}$.*

The argument developed in the proof of Theorem 77 can be applied here without essential modifications. Thus we find that the sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ are disjoint and cover $\mathfrak{R}$; that every function in $\mathfrak{N}$ is constant on each set $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$; that every set in $\mathcal{G}_{\mathfrak{N}}$ is the union of the sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ which it contains; and that

$\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, as the intersection of all the closed sets $\mathfrak{R}(f=f(\mathfrak{r}))$ where $f\epsilon\mathfrak{N}$, is itself a closed set. Moreover, we see that, if $\mathfrak{G}_{\mathfrak{N}}^{(k)}=\mathfrak{R}(\alpha_k<f_k<\beta_k)$, $f_k\epsilon\mathfrak{N}$, for $k=1,\cdots,n$, then the intersection $\mathfrak{H}_{\mathfrak{N}}=\mathfrak{G}_{\mathfrak{N}}^{(1)}\cdots\mathfrak{G}_{\mathfrak{N}}^{(n)}$ is the union of all the sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ which it contains: for $\mathfrak{r}\epsilon\mathfrak{H}_{\mathfrak{N}}$ implies $\alpha_k<f_k(\mathfrak{r})<\beta_k$ for $k=1,\cdots,n$ and hence $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{G}_{\mathfrak{N}}^{(k)}$ for $k=1,\cdots,n$. Thus the neighborhood-system described in the theorem can be imposed upon the family $\mathcal{X}_{\mathfrak{N}}$; and the resulting space $\mathfrak{R}_{\mathfrak{N}}$ is a T_1-space. As in the proof of Theorem 77, we see next that the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ from $\mathfrak{R}$ to $\mathfrak{R}_{\mathfrak{N}}$ is univocal and continuous; and that it carries each function f in $\mathfrak{N}$ into a bounded continuous function f^* in $\mathfrak{R}_{\mathfrak{N}}$ given by $f^*(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=f(\mathfrak{r})$. It is now easily seen that $\mathfrak{R}_{\mathfrak{N}}$ is a CR-space: for the open set in $\mathfrak{R}_{\mathfrak{N}}$ specified by the relation $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{H}_{\mathfrak{N}}$, where $\mathfrak{H}_{\mathfrak{N}}=\mathfrak{R}(\alpha_1<f_1<\beta_1)\cdots\mathfrak{R}(\alpha_n<f_n<\beta_n)$ and $f_k\epsilon\mathfrak{N}$ for $k=1,\cdots,n$, coincides with the set

$$\mathfrak{R}_{\mathfrak{N}}(\alpha_1<f_1^*<\beta_1)\cdots\mathfrak{R}_{\mathfrak{N}}(\alpha_n<f_n^*<\beta_n)=\mathfrak{R}_{\mathfrak{N}}(0<g_{\mathfrak{N}}<\gamma)$$

where $g_{\mathfrak{N}}$ is the bounded continuous real function in $\mathfrak{R}_{\mathfrak{N}}$ defined by the formula

$$g_{\mathfrak{N}}=\min_{k=1,\cdots,n}\left[(f_k^*-\alpha_k)(\beta_k-f_k^*)\right]$$

and $\gamma>\|g_{\mathfrak{N}}\|$. As an H-space which is a continuous image of the bicompact H-space $\mathfrak{R}$, the space $\mathfrak{R}_{\mathfrak{N}}$ must also be bicompact. In case $\mathfrak{N}=\mathfrak{M}$, the construction of $\mathfrak{R}_{\mathfrak{M}}$ is identical with that described in Theorem 77, so that $\mathfrak{R}_{\mathfrak{M}}$ coincides with $\mathfrak{R}^*$ and is topologically equivalent to $\mathfrak{R}$. If $\mathfrak{N}_1\supset\mathfrak{N}_2$, the relations $\mathcal{G}_{\mathfrak{N}_1}\supset\mathcal{G}_{\mathfrak{N}_2}$, $\mathfrak{X}_{\mathfrak{N}_1}(\mathfrak{r})\subset\mathfrak{X}_{\mathfrak{N}_2}(\mathfrak{r})$ are obviously valid. It follows immediately that the correspondence $\mathfrak{X}_{\mathfrak{N}_1}(\mathfrak{r})\rightarrow\mathfrak{X}_{\mathfrak{N}_2}(\mathfrak{r})$ from $\mathfrak{R}_{\mathfrak{N}_1}$ to $\mathfrak{R}_{\mathfrak{N}_2}$ is univocal and continuous; and that $\mathfrak{R}_{\mathfrak{N}_2}$ is a continuous image of $\mathfrak{R}_{\mathfrak{N}_1}$. By inspection of the construction of the analytical subring $\mathfrak{N}^*$ generated by $\mathfrak{N}$, it is evident that every function in $\mathfrak{N}^*$ is constant on each set $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ and hence that $\mathfrak{X}_{\mathfrak{N}^*}(\mathfrak{r})\supset\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$. The relation $\mathfrak{N}^*\supset\mathfrak{N}$ implies that $\mathfrak{R}_{\mathfrak{N}}$ is a continuous image of $\mathfrak{R}_{\mathfrak{N}^*}$ under the correspondence $\mathfrak{X}_{\mathfrak{N}^*}(\mathfrak{r})\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ where $\mathfrak{X}_{\mathfrak{N}^*}(\mathfrak{r})\subset\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$. We see therefore that $\mathfrak{X}_{\mathfrak{N}^*}(\mathfrak{r})=\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ and that the indicated correspondence is biunivocal. An elementary proposition now shows that the bicompact H-spaces $\mathfrak{R}_{\mathfrak{N}}$ and $\mathfrak{R}_{\mathfrak{N}^*}$ are topologically equivalent.†

In order to obtain a deeper insight into the subject matter of Theorem 81, we shall next prove a generalization of the Weierstrass approximation theorem.

THEOREM 82. *In order that $\mathfrak{R}_{\mathfrak{N}}$ be topologically equivalent to $\mathfrak{R}_{\mathfrak{M}}$ and to $\mathfrak{R}$ by virtue of the correspondences* $\mathfrak{r}\rightarrow\mathfrak{X}_{\mathfrak{M}}(\mathfrak{r})\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, *it is necessary and sufficient that* $\mathfrak{N}^*=\mathfrak{M}$.

† AH, p. 95, Satz III.

The usual statement of the Weierstrass approximation theorem can be broken down into two propositions: (1) every function continuous in a closed interval of R can be uniformly approximated by a "polynomial" constructed from the function $f(\alpha)=\alpha$ and the real numbers by the algebraic operations and the formation of absolute values; (2) the function $f(\alpha)=|\alpha|$ can be uniformly approximated by a polynomial $p(\alpha)$ where $p(0)=0$. It is obvious that $\mathfrak{N}^*$ can be constructed from $\mathfrak{N}$ by taking the uniform limits of "polynomials" constructed from $\mathfrak{N}$ and R by the application of the algebraic operations and the formation of absolute values. Hence, if we restrict ourselves to this characterization of $\mathfrak{N}^*$, the equation $\mathfrak{N}^*=\mathfrak{M}$ may be regarded as a generalization of part (1) of the Weierstrass approximation theorem. The proof which we shall give below will be a proof of this partial generalization; it will be in particular a proof of (1). In order to have a proof of a complete generalization of the Weierstrass approximation theorem, we must eliminate the use of absolute values from the construction of $\mathfrak{N}^*$. We can do this, as we have already seen in the discussion of Theorem 75, if and only if we use part (2) of the Weierstrass approximation theorem. Since our proof of the generalization of part (1) will be an algebraico-topological proof, we may regard (2) as the "analytical kernel" of the Weierstrass approximation theorem for general topological spaces.

From Theorem 81 we already know that $\mathfrak{R}_{\mathfrak{N}}$ and $\mathfrak{R}_{\mathfrak{N}^*}$ are topologically equivalent. Hence there is no loss of generality in assuming that $\mathfrak{N}=\mathfrak{N}^*$, in other words, that $\mathfrak{N}$ is an analytical subring. From Theorem 81 it is then known that the relation $\mathfrak{N}=\mathfrak{M}$ implies the topological equivalence of $\mathfrak{R}_{\mathfrak{N}}$, $\mathfrak{R}_{\mathfrak{M}}$, and $\mathfrak{R}$. Hence we have only to prove that the topological equivalence of $\mathfrak{R}_{\mathfrak{N}}$ and $\mathfrak{R}_{\mathfrak{M}}$ implies $\mathfrak{N}=\mathfrak{M}$ when $\mathfrak{N}$ is an analytical subring of $\mathfrak{M}$.

If f is an arbitrary function in $\mathfrak{M}$ and ϵ is an arbitrary positive number, we shall construct a function g in $\mathfrak{N}$ with the properties $f\leqq g<f+\epsilon$. Since $\mathfrak{N}$ is closed in $\mathfrak{M}$, the resulting inequality $\|f-g\|<\epsilon$ leads to the conclusion that $\mathfrak{N}=\mathfrak{M}$.

As a first step in this construction we show that, if α and β are real numbers satisfying the inequality $\alpha<\beta$, then there exists a function $g_{\alpha\beta}$ in $\mathfrak{N}$ which satisfies the inequality $f\leqq g_{\alpha\beta}$ in $\mathfrak{R}$ and the inequality $g_{\alpha\beta}\leqq\beta+\epsilon/2$ in the closed set $\mathfrak{F}=\mathfrak{R}(\alpha\leqq f\leqq\beta)$. If $\beta+\epsilon/2\geqq\|f\|$ or if $\mathfrak{F}$ is void, we may obviously take $g_{\alpha\beta}$ as the constant $\|f\|$; and if $\mathfrak{G}=\mathfrak{R}(\alpha-\epsilon/2<f<\beta+\epsilon/2)$ coincides with $\mathfrak{R}$, we may take $g_{\alpha\beta}$ as the constant $\beta+\epsilon/2$. Hence we may confine our attention to the case where $\beta+\epsilon/2<\|f\|$, $\mathfrak{F}\neq 0$, and $\mathfrak{G}\neq\mathfrak{R}$. Since $\mathfrak{R}_{\mathfrak{N}}$ is topologically equivalent to $\mathfrak{R}_{\mathfrak{M}}$ and to $\mathfrak{R}$ by hypothesis, and since $\mathfrak{G}$ contains $\mathfrak{F}$, each point $\mathfrak{r}$ in $\mathfrak{F}$ determines an open set $\mathfrak{G}_{\mathfrak{N}}(\mathfrak{r})=\mathfrak{R}(\gamma<g<\delta)$, $g\epsilon\mathfrak{N}$, such that $\mathfrak{r}\epsilon\mathfrak{G}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{G}$. Now the function $h=|g-g(\mathfrak{r})|/\eta$, where $0<\eta<\min\ (g(\mathfrak{r})-\gamma,\ \delta-g(\mathfrak{r}))$, be-

longs to $\mathfrak{N}$; and the sets $\mathfrak{H}_{\mathfrak{N}}(\mathfrak{r})=\mathfrak{N}(h<1)$, $\mathfrak{F}_{\mathfrak{N}}(\mathfrak{r})=\mathfrak{N}(h\leqq 1)$ are respectively open and closed in $\mathfrak{N}$ and satisfy the relations $\mathfrak{r}\epsilon\mathfrak{H}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{F}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{G}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{G}$. As a closed set in the bicompact H-space $\mathfrak{N}$, the subspace $\mathfrak{F}$ is bicompact. Hence there exist points $\mathfrak{r}_1, \cdots, \mathfrak{r}_n$ in $\mathfrak{F}$ such that

$$\mathfrak{F}\subset\mathfrak{H}_{\mathfrak{N}}(\mathfrak{r}_1)\cup\cdots\cup\mathfrak{H}_{\mathfrak{N}}(\mathfrak{r}_n)\subset\mathfrak{F}_{\mathfrak{N}}(\mathfrak{r}_1)\cup\cdots\cup\mathfrak{F}_{\mathfrak{N}}(\mathfrak{r}_n)\subset\mathfrak{G}.$$

If the corresponding functions are $h_1, \cdots, h_n$ respectively, the function $h=\min(h_1, \cdots, h_n)$ belongs to $\mathfrak{N}$ and has the properties

$$\mathfrak{H}_{\mathfrak{N}} = \mathfrak{N}(h < 1) = \mathfrak{H}_{\mathfrak{N}}(\mathfrak{r}_1)\cup\cdots\cup\mathfrak{H}_{\mathfrak{N}}(\mathfrak{r}_n),$$
$$\mathfrak{F}_{\mathfrak{N}} = \mathfrak{N}(h \leqq 1) = \mathfrak{F}_{\mathfrak{N}}(\mathfrak{r}_1)\cup\cdots\cup\mathfrak{F}_{\mathfrak{N}}(\mathfrak{r}_n),$$
$$\mathfrak{F}\subset\mathfrak{H}_{\mathfrak{N}}\subset\mathfrak{F}_{\mathfrak{N}}\subset\mathfrak{G}.$$

Since $\mathfrak{G}'$ is a non-void closed subset of the bicompact space $\mathfrak{N}$, the function h has on this set a greatest lower bound ξ attained at some point $\mathfrak{r}$ in $\mathfrak{G}'$; and the relations $\mathfrak{G}'\subset\mathfrak{F}'_{\mathfrak{N}}=\mathfrak{N}(h>1)$ show that $\xi=f(\mathfrak{r})>1$. We can now define the desired function $g_{\alpha\beta}$ by the formula

$$g_{\alpha\beta} = \{\|f\| - (\beta+\epsilon/2)\}\{\min[\xi, \max(h, 1)] - 1\}/\{\xi - 1\} + (\beta+\epsilon/2).$$

It is evident that $g_{\alpha\beta}$ is in $\mathfrak{N}$. The formula, together with the relations $\beta+\epsilon/2<\|f\|$ and $\xi>1$, makes it plain that $g_{\alpha\beta}$ satisfies the inequality $\beta+\epsilon/2\leqq g_{\alpha\beta}\leqq\|f\|$. Since $\mathfrak{G}'\subset\mathfrak{N}(h\geqq\xi)$ in accordance with our determination of ξ, we see that $g_{\alpha\beta}=\|f\|\geqq f$ in $\mathfrak{G}'$; on the other hand, we have $g_{\alpha\beta}\geqq\beta+\epsilon/2>f$ in $\mathfrak{G}$. Hence the inequality $f\leqq g_{\alpha\beta}$ holds at every point in $\mathfrak{N}$. Since $\mathfrak{F}\subset\mathfrak{F}_{\mathfrak{N}}=\mathfrak{N}(h\leqq 1)$ we see that $g_{\alpha\beta}=\beta+\epsilon/2$ in $\mathfrak{F}$. Thus the function $g_{\alpha\beta}$ has all the properties required.

With the aid of the functions $g_{\alpha\beta}$, we can now construct the desired function g. We choose a finite number of open intervals (α, β) of positive length $\beta-\alpha$ not greater than $\epsilon/2$ which cover the range of the given function f in $\mathfrak{M}$. Let the functions $g_{\alpha\beta}$ corresponding to these intervals be denoted by $g_1, \cdots, g_n$ respectively. The function $g=\min(g_1, \cdots, g_n)$ then belongs to $\mathfrak{N}$. It obviously has the property $g\geqq f$ since $g_k\geqq f$ for $k=1, \cdots, n$. If $\mathfrak{r}$ is any point in $\mathfrak{N}$, then $f(\mathfrak{r})$ belongs to at least one interval (α, β) among those chosen above; and the relations $\alpha<f(\mathfrak{r})<\beta$, $0<\beta-\alpha\leqq\epsilon/2$ show that $\beta<f(\mathfrak{r})+\epsilon/2$. Hence we have $g(\mathfrak{r})\leqq g_{\alpha\beta}(\mathfrak{r})\leqq\beta+\epsilon/2<f(\mathfrak{r})+\epsilon$; and the inequality $g<f+\epsilon$ holds throughout $\mathfrak{N}$. With this the proof is completed.

We may point out in detail that this proof establishes part (1) of the Weierstrass approximation theorem. If $\mathfrak{N}$ is a closed subinterval of R, then the function f given by $f(\rho)=\rho$ has the property that the sets $\mathfrak{N}(\alpha<f<\beta)$ constitute a basis for $\mathfrak{N}$. Hence if we take $\mathfrak{N}$ as the subclass of $\mathfrak{M}$ consisting

of the function f alone, $\mathfrak{R}_{\mathfrak{R}}$ is topologically equivalent to $\mathfrak{R}$ and $\mathfrak{R}^* = \mathfrak{M}$, as we wished to prove.

Before applying the generalized approximation theorem, we shall generalize a theorem established by Banach in the case of separable bicompact H-spaces (compact metric spaces). Our proof is necessarily somewhat different from the proof of Banach.†

THEOREM 83. *If $\mathfrak{R}$ and $\mathfrak{R}^*$ are bicompact H-spaces and if $\mathfrak{M}$ and $\mathfrak{M}^*$ are the corresponding function-rings, then the existence of an isometric correspondence $f \to Uf = f^*$ between $\mathfrak{M}$ and $\mathfrak{M}^*$ is equivalent to the existence of a topological equivalence $\mathfrak{r} \to \rho(\mathfrak{r}) = \mathfrak{r}^*$ between $\mathfrak{R}$ and $\mathfrak{R}^*$, the two correspondences being connected by the relations*

$$f(\mathfrak{r}) = \phi^*(\mathfrak{r}^*)[f^*(\mathfrak{r}^*) - \theta^*(\mathfrak{r}^*)], \qquad Uf = f^*, \qquad \rho(\mathfrak{r}) = \mathfrak{r}^*,$$
$$U0 = \theta^*, \qquad U1 = \phi^*,$$

where $|\phi^*| = 1$. If the relations $U0 = 0$, $U1 = 1$ or, equivalently, the relations $\theta^* = 0$, $\phi^* = 1$ are satisfied, then the correspondence $f \to Uf = f^*$ is an analytical isomorphism between $\mathfrak{M}$ and $\mathfrak{M}^*$.

By an isometric correspondence U is meant one with the property $\|Uf - Ug\| = \|f - g\|$. It is evident that such a correspondence is biunivocal and has an isometric inverse U^{-1}.

If ρ, ϕ^*, and θ^* are given arbitrarily, we define U by the equation $f^* = Uf = f(\rho(\mathfrak{r}))/\phi^*(\rho(\mathfrak{r})) + \theta^*(\rho(\mathfrak{r}))$. It is easily verified that U carries $\mathfrak{M}$ isometrically into $\mathfrak{M}^*$. With the special choice $\phi^* = 1$, $\theta^* = 0$, it is evident that U determines an analytical isomorphism between $\mathfrak{M}$ and $\mathfrak{M}^*$. We may remark that if U defines such an isomorphism the necessary relations $U0 = 0$, $U1 = 1$ imply $\theta^* = 0$, $\phi^* = 1$.

When U is given, we define $\theta^* = U0$ and determine a new correspondence V by the relations $f \to Vf = Uf - \theta^* = f^* - \theta^*$. It is evident that V carries $\mathfrak{M}$ isometrically into $\mathfrak{M}^*$ and that V has the additional property $V0 = 0$. A theorem of Mazur and Ulam now shows that V is a linear correspondence, satisfying the relation $V(\alpha f + \beta g) = \alpha Vf + \beta Vg$.‡

We now construct the topological equivalence $\rho(\mathfrak{r})$ in terms of the correspondence V. If $\mathfrak{r}$ is any point in $\mathfrak{R}$ we denote by $\mathfrak{M}(\mathfrak{r})$ the class of all functions f in $\mathfrak{M}$ which satisfy the equation $|f(\mathfrak{r})| = \|f\|$. It is evident that $\mathfrak{M}(\mathfrak{r}) \supset R$. Also, if $f_1, \cdots, f_n$ are in $\mathfrak{M}(\mathfrak{r})$, the function $g = \sum_{\nu=1}^{\nu=n} f_\nu \operatorname{sgn} f_\nu(\mathfrak{r})$ belongs to $\mathfrak{M}(\mathfrak{r})$ and satisfies the relation $\|g\| = \sum_{\nu=1}^{\nu=n} \|f_\nu\|$, as we infer from the inequalities

† See Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 170, Théorème 3.

‡ A proof is given by Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, pp. 166–168.

$$\|g\| \leqq \sum_{\nu=1}^{\nu=n} \|f_\nu\| = \sum_{\nu=1}^{\nu=n} |f_\nu(\mathfrak{r})| = g(\mathfrak{r}) \leqq \|g\|.$$

We can now show that the closed sets $\mathfrak{F}^* = \mathfrak{R}^*(|f^*(\mathfrak{r}^*)| = \|f^*\|)$, where $f^* = Vf$ and $f \epsilon \mathfrak{M}(\mathfrak{r})$, have a non-void intersection in $\mathfrak{R}^*$. Since $\mathfrak{R}^*$ is bicompact, it suffices to prove that the intersection of a finite number of sets $\mathfrak{F}^*$ is non-void. If $\mathfrak{F}_1^*, \cdots, \mathfrak{F}_n^*$ are such sets corresponding to the respective functions $f_1, \cdots, f_n$ in $\mathfrak{M}(\mathfrak{r})$, we consider the associated function g defined above and its correspondent

$$g^* = Vg = \sum_{\nu=1}^{\nu=n} V(f_\nu \operatorname{sgn} f_\nu(\mathfrak{r})) = \sum_{\nu=1}^{\nu=n} (Vf_\nu) \operatorname{sgn} f_\nu(\mathfrak{r}) = \sum_{\nu=1}^{\nu=n} f_\nu^* \operatorname{sgn} f_\nu(\mathfrak{r}).$$

Since g^* assumes its greatest lower and least upper bounds on the bicompact space $\mathfrak{R}^*$, there exists a point $\mathfrak{p}^*$ in $\mathfrak{R}^*$ such that $|g^*(\mathfrak{p}^*)| = \|g^*\|$. We now observe the relations

$$\|g^*\| = |g^*(\mathfrak{p}^*)| \leqq \sum_{\nu=1}^{\nu=n} |f_\nu^*(\mathfrak{p}^*)| \leqq \sum_{\nu=1}^{\nu=n} \|f_\nu^*\| = \sum_{\nu=1}^{\nu=n} \|f_\nu\| = \|g\| = \|g^*\|$$

and the relations

$$\sum_{\nu=1}^{\nu=n} |f_\nu^*(\mathfrak{p}^*)| = \sum_{\nu=1}^{\nu=n} \|f_\nu^*\|, \qquad |f_k^*(\mathfrak{p}^*)| \leqq \|f_k^*\|, \qquad k = 1, \cdots, n.$$

The latter show that we must have $|f_k^*(\mathfrak{p}^*)| = \|f_k^*\|$ for $k = 1, \cdots, n$. Hence the point $\mathfrak{p}^*$ is common to $\mathfrak{F}_1^*, \cdots, \mathfrak{F}_n^*$ as we wished to prove. We now let $\mathfrak{r}^*$ be a point common to all the sets $\mathfrak{F}^*$. It is evident that $f \epsilon \mathfrak{M}(\mathfrak{r})$ and $f^* = Vf$ imply $f^* \epsilon \mathfrak{M}^*(\mathfrak{r}^*)$; in other words, that V carries $\mathfrak{M}(\mathfrak{r})$ into a subclass of $\mathfrak{M}^*(\mathfrak{r}^*)$. By symmetry, the inverse correspondence V^{-1} carries $\mathfrak{M}^*(\mathfrak{r}^*)$ into a subclass of some class $\mathfrak{M}(\mathfrak{p})$. The inclusion relation $\mathfrak{M}(\mathfrak{p}) \supset \mathfrak{M}(\mathfrak{r})$ is now obvious; and it implies that $\mathfrak{p} = \mathfrak{r}$. In fact, if $\mathfrak{p} \neq \mathfrak{r}$, there exists a function f in $\mathfrak{M}$ such that $f(\mathfrak{p}) = 0$, $f(\mathfrak{r}) = 1$, $0 \leqq f \leqq 1$; and the equation $\|f\| = 1$ then shows that $f \epsilon \mathfrak{M}(\mathfrak{r})$, $f \not\epsilon \mathfrak{M}(\mathfrak{p})$. We can now infer that V carries $\mathfrak{M}(\mathfrak{r})$ into $\mathfrak{M}^*(\mathfrak{r}^*)$; and, further, that V carries the family of all classes $\mathfrak{M}(\mathfrak{r})$ biunivocally into the family of all classes $\mathfrak{M}^*(\mathfrak{r}^*)$. This correspondence between the classes $\mathfrak{M}(\mathfrak{r})$ and $\mathfrak{M}^*(\mathfrak{r}^*)$ determines a biunivocal correspondence $\mathfrak{r} \longleftrightarrow \mathfrak{r}^* = \rho(\mathfrak{r})$ between $\mathfrak{R}$ and $\mathfrak{R}^*$. We have to prove that the latter correspondence is a topological equivalence. From the construction of $\rho(\mathfrak{r})$, it is seen that the sets $\mathfrak{R}(|f| = \|f\|)$ and $\mathfrak{R}^*(|f^*| = \|f^*\|)$ correspond under the correspondence $\mathfrak{r} \longleftrightarrow \mathfrak{r}^* = \rho(\mathfrak{r})$ when f and f^* are connected by the relation $f^* = Vf$: for these two sets are specified by the relations $f \epsilon \mathfrak{M}(\mathfrak{r})$ and $f^* \epsilon \mathfrak{M}^*(\mathfrak{r}^*)$ respectively. Accordingly the complementary sets $\mathfrak{R}(|f| < \|f\|)$ and $\mathfrak{R}^*(|f^*| < \|f^*\|)$ correspond likewise.

Thus it is sufficient for us to prove that the latter sets constitute bases for the respective spaces $\mathfrak{R}$ and $\mathfrak{R}^*$. Since the same discussion applies to both $\mathfrak{R}$ and $\mathfrak{R}^*$, we may consider the space $\mathfrak{R}$ alone. Since $\mathfrak{R}$ is a bicompact H-space, it is a CR-space; and the sets $\mathfrak{R}(\alpha<g<\beta)$, $g\epsilon\mathfrak{M}$, constitute a basis for $\mathfrak{R}$. If $\mathfrak{R}$ has more than one point, we may discard the sets $\mathfrak{R}(\alpha<g<\beta)$ which coincide with $\mathfrak{R}$ or are void, and still have a basis for $\mathfrak{R}$. If $\mathfrak{R}(\alpha<g<\beta)$ is one of the sets retained, we introduce the function

$$f = 1 - \max\,[0, (g - \alpha)(\beta - g)]/\|(g - \alpha)(\beta - g)\|$$

and show that $\mathfrak{R}(\alpha<g<\beta)=\mathfrak{R}(|f|<\|f\|)$. It is evident that $0\leqq f\leqq 1$. If $\mathfrak{r}$ is in the complement of $\mathfrak{R}(\alpha<g<\beta)$, we have $(g(\mathfrak{r})-\alpha)(\beta-g(\mathfrak{r}))\leqq 0$ and $f=1$. We infer that $\|f\|=1$. On the other hand, if $\mathfrak{r}\epsilon\mathfrak{R}(\alpha<g<\beta)$, we have $(g(\mathfrak{r})-\alpha)(\beta-g(\mathfrak{r}))>0$ and $f<1$. Thus we have $\mathfrak{R}(\alpha<g<\beta)=\mathfrak{R}(|f|<\|f\|)$, as we wished to prove. It is evident that f is in $\mathfrak{M}$. Thus we have proved that, unless $\mathfrak{R}$ and $\mathfrak{R}^*$ are one-element spaces, the correspondence $\mathfrak{r}\longleftrightarrow\mathfrak{r}^*=\rho(\mathfrak{r})$ is a topological equivalence; and the exceptional case is trivial.

We now define ϕ^* as the function $V1$ in $\mathfrak{M}^*$. Since $\|\phi^*\|=\|f\|=1$ and $\mathfrak{R}(|f|=\|f\|)=\mathfrak{R}$ for $f=1$, we conclude that $\mathfrak{R}^*(|\phi^*|=\|\phi^*\|)=\mathfrak{R}^*$ and hence that $|\phi^*|=1$. The correspondence W defined by $Wf=\phi^*Vf$ therefore has the properties $|Wf|=|Vf|$, $\|Wf\|=\|Vf\|=\|f\|$, $W0=0$, $W1=\phi^*\phi^*=1$, and $W(\alpha f+\beta g)=\alpha Wf+\beta Wg$. Since the first two of these properties imply $\mathfrak{R}^*(|Vf|=\|Vf\|)=\mathfrak{R}^*(|Wf|=\|Wf\|)$, we see that the construction of the preceding paragraph leads to the same topological equivalence ρ if we start with W rather than with V.

In terms of W the relation between U, ϕ^*, θ^*, and ρ which we wish to establish assumes the equivalent but simpler form $f(\mathfrak{r})=f^*(\mathfrak{r}^*)$ where $f^*=Wf$ and $\mathfrak{r}^*=\rho(\mathfrak{r})$. As a first step in proving this relation, we show that $f\geqq 0$ implies $Wf\geqq 0$. If α and β are the minimum and maximum, respectively, of the function f, the relations $0\leqq\alpha\leqq\beta$ imply that the function $g=\beta-f\geqq 0$ has the number $\beta-\alpha\geqq 0$ as its maximum. Hence $\|g\|=\beta-\alpha$. If we now write $f^*=Wf=W\beta+Wf-W\beta=\beta-Wg\geqq\beta-\|Wg\|=\beta-\|g\|=\alpha\geqq 0$, we obtain the desired result. As a second step, we prove that $W|f|=|Wf|$. Since $|f|-f\geqq 0$, we have $W|f|-Wf=W(|f|-f)\geqq 0$ and hence $W|f|\geqq Wf$. Similarly $|f|+f\geqq 0$ implies $-W|f|\leqq Wf$. We therefore conclude that $W|f|\geqq|Wf|$. By symmetry, $W^{-1}|Wf|\geqq|W^{-1}Wf|=|f|$. Since $W^{-1}|Wf|-|f|\geqq 0$, we have $|Wf|-W|f|=W(W^{-1}|Wf|-|f|)\geqq 0$ and hence $|Wf|\geqq W|f|$. Combining this inequality with the one obtained above, we conclude that $|Wf|=W|f|$. We are now in a position to complete our proof. Let α be the value of f at a fixed point $\mathfrak{r}$ in $\mathfrak{R}$ and let β be the maximum of the function $|f-\alpha|$. Then the function $g=\beta-|f-\alpha|\geqq 0$ belongs to $\mathfrak{M}$ and has a maximum β at the point $\mathfrak{r}$.

Hence $\|g\|=\beta$. Since $\mathfrak{r}\epsilon\mathfrak{R}(|g|=\|g\|)$, we see that $\mathfrak{r}^*=\rho(\mathfrak{r})$ belongs to the set $\mathfrak{R}(|g^*|=\|g^*\|)$ where $g^*=Wg$. Now $g^*=Wg=W(\beta-|f-\alpha|)=\beta-W|f-\alpha|$ $=\beta-|Wf-W\alpha|=\beta-|f^*-\alpha|$, $|g^*|=|Wg|=W|g|=Wg=g^*$, and $\|g^*\|$ $=\|g\|=\beta$. Hence we see that $\beta-|f^*(\mathfrak{r}^*)-\alpha|=g^*(\mathfrak{r}^*)=|g^*(\mathfrak{r}^*)|=\|g^*\|=\beta$, $f^*(\mathfrak{r}^*)=\alpha=f(\mathfrak{r})$. This completes the demonstration.

We can now return to the analysis of the results of Theorem 81, obtaining the following additional information.

THEOREM 84. *The correspondence* $\mathfrak{r}\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ *from* $\mathfrak{R}$ *to* $\mathfrak{R}_{\mathfrak{N}}$ *described in Theorem* 81 *induces an analytical isomorphism between* $\mathfrak{N}^*$ *and the function-ring* $\mathfrak{M}_{\mathfrak{N}}$ *for* $\mathfrak{R}_{\mathfrak{N}}$. *No function which belongs to* $\mathfrak{M}$ *but not to* $\mathfrak{N}^*$ *is carried by this correspondence into a single-valued function in* $\mathfrak{R}_{\mathfrak{N}}$.

The correspondence $\mathfrak{r}\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ defines a function $f_{\mathfrak{N}}$ in $\mathfrak{R}_{\mathfrak{N}}$ through the relation $f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=f(\mathfrak{r})$. If f belongs to $\mathfrak{N}$ or $\mathfrak{N}^*$, as we have already noted, then $f_{\mathfrak{N}}$ is single-valued, bounded and real. If f is in $\mathfrak{N}$, then $f_{\mathfrak{N}}$ is also cont nuous. Thus the induced correspondence $f\rightarrow f_{\mathfrak{N}}$ carries $\mathfrak{N}$ into a subclass $\mathfrak{P}\mathrm{i}_{\mathfrak{N}}$ of the function ring $\mathfrak{M}_{\mathfrak{N}}$ for $\mathfrak{R}_{\mathfrak{N}}$ in such a way that $|f|\rightarrow|f_{\mathfrak{N}}|$ and $\|f\|=\|f_{\mathfrak{N}}\|$. It is obvious that the space $(\mathfrak{R}_{\mathfrak{N}})_{\mathfrak{P}_{\mathfrak{N}}}$ is topologically equivalent to $\mathfrak{R}_{\mathfrak{N}}$. Theorem 82 therefore shows that the analytical subring $(\mathfrak{P}_{\mathfrak{N}})^*$ generated in $\mathfrak{M}_{\mathfrak{N}}$ by $\mathfrak{P}_{\mathfrak{N}}$ coincides with $\mathfrak{M}_{\mathfrak{N}}$. Obviously the correspondence $f\rightarrow f_{\mathfrak{N}}$ can be extended analytically from the classes $\mathfrak{N}$, $\mathfrak{P}_{\mathfrak{N}}$ to the corresponding classes $\mathfrak{N}^*$, $(\mathfrak{P}_{\mathfrak{N}})^*=\mathfrak{M}_{\mathfrak{N}}$. The extended correspondence is then seen to be an analytical isomorphism between $\mathfrak{N}^*$ and $\mathfrak{M}_{\mathfrak{N}}$; it coincides with the correspondence $f\rightarrow f_{\mathfrak{N}}$ defined above for $f\epsilon\mathfrak{N}^*$ before we had shown that $f_{\mathfrak{N}}\epsilon\mathfrak{M}_{\mathfrak{N}}$. If f is not in $\mathfrak{N}^*$, we wish to show that the corresponding function $f_{\mathfrak{N}}$ defined by $f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=f(\mathfrak{r})$ is not single-valued. To this end we consider the class $\mathfrak{P}$ of all functions in $\mathfrak{M}$ which are constant on each set $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, $\mathfrak{r}\epsilon\mathfrak{R}$. Since $\mathfrak{P}\supset\mathfrak{N}^*\supset\mathfrak{N}$, the space $\mathfrak{R}_{\mathfrak{N}}$ is a continuous image of $\mathfrak{R}_{\mathfrak{P}}$ by virtue of the correspondence $\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r})\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ as we proved in Theorem 81. The definition of $\mathfrak{P}$ shows that $\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r})=\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, so that this correspondence is biunivocal. Since $\mathfrak{R}_{\mathfrak{P}}$ and $\mathfrak{R}_{\mathfrak{N}}$ are bicompact H-spaces, it follows that they must be topologically equivalent.† If $f\epsilon\mathfrak{P}$, the function $f_{\mathfrak{P}}$ in $\mathfrak{R}_{\mathfrak{P}}$ defined by $f_{\mathfrak{P}}(\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r}))=f(\mathfrak{r})$ is a bounded continuous real function in accordance with Theorem 81. The correspondence $\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r})\rightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ therefore carries $f_{\mathfrak{P}}$ into the bounded continuous real function $f_{\mathfrak{N}}$ defined by $f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))$ $=f_{\mathfrak{P}}(\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r}))=f(\mathfrak{r})$. Since $f_{\mathfrak{N}}$ is in $\mathfrak{M}_{\mathfrak{N}}$, there exists a function g in $\mathfrak{N}^*$ such that $g_{\mathfrak{N}}=f_{\mathfrak{N}}$, by virtue of the results established above. Hence we have $f(\mathfrak{r})=f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=g_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=g(\mathfrak{r})$ throughout $\mathfrak{R}$; and we conclude that $f\epsilon\mathfrak{N}^*$. We have thus shown that $\mathfrak{N}^*\supset\mathfrak{P}$. It follows that $\mathfrak{P}=\mathfrak{N}^*$. We therefore see

† AH, p. 95, Satz III.

that, if f is not in $\mathfrak{N}^*$, then f is not constant on every set $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ and that the corresponding function $f_{\mathfrak{N}}$ is not single-valued.

With Theorems 80–84 we have obtained enough information to characterize completely the algebraico-topological structure of the function-rings for bicompact H-spaces. We state first the results for the theory of closed ideals.

THEOREM 85. *Between the closed ideals in the function-ring $\mathfrak{M}$ for a bicompact H-space $\mathfrak{R}$ and the closed subsets of $\mathfrak{R}$, there exists a biunivocal correspondence such that a closed ideal $\mathfrak{N}$ consists of all functions in $\mathfrak{M}$ which vanish on the corresponding closed set $\mathfrak{F}$. The quotient-ring $\mathfrak{M}/\mathfrak{N}$, where $\mathfrak{N}$ is a closed ideal distinct from $\mathfrak{M}$, is isomorphic to the function-ring of the corresponding set $\mathfrak{F}$. Every product of divisorless ideals is a closed ideal; and every closed ideal distinct from $\mathfrak{M}$ is the product of all the divisorless ideals which contain it.*

We denote by $\mathfrak{G}$ the set of all points where some function f in a given closed ideal $\mathfrak{N}$ does not vanish; and by $\mathfrak{F}$ the set of all points where every function f in $\mathfrak{N}$ vanishes. It is evident that $\mathfrak{G}$ and $\mathfrak{F}$ are complements of one another. We see also that $\mathfrak{G}$ is open: for if $\mathfrak{r}\epsilon\mathfrak{G}$ and $f(\mathfrak{r})=\alpha\neq 0, f\epsilon\mathfrak{N}$, the open set $\mathfrak{R}(\alpha-\epsilon<f<\alpha+\epsilon)$ contains $\mathfrak{r}$ and is contained in $\mathfrak{G}$ whenever $0<\epsilon<|\alpha|$. Accordingly, $\mathfrak{F}$ must be a closed set. At this point, it is convenient to show that $\mathfrak{F}=0$ implies $\mathfrak{N}=\mathfrak{M}$. Since $\mathfrak{F}=0$ implies $\mathfrak{G}=\mathfrak{R}$, the open sets $\mathfrak{R}(\alpha<f<\beta)$, where $f\epsilon\mathfrak{N}$ and 0 is outside the closed interval $[\alpha, \beta]$, cover $\mathfrak{R}$. Since $\mathfrak{R}$ is bicompact, we can select a finite number of these sets which also cover $\mathfrak{R}$. If the corresponding functions in $\mathfrak{N}$ are $f_1, \cdots, f_n$, we see that the function $f=|f_1|+\cdots+|f_n|$ has a positive lower bound in $\mathfrak{R}$. If f were in $\mathfrak{N}$, we could then conclude, since $\mathfrak{N}$ is an ideal, that $1=f\cdot(1/f)\epsilon\mathfrak{N}$ and hence that $\mathfrak{N}=\mathfrak{M}$. To prove that f actually belongs to $\mathfrak{N}$, it is evidently sufficient to show that $\mathfrak{N}$ contains $|f|$ together with f. Since $|f|$ can be uniformly approximated by polynomials of the form

$$\sum_{\nu=1}^{\nu=n}\alpha_\nu f^\nu = f\left(\sum_{\nu=0}^{\nu=n-1}\alpha_{\nu+1}f^\nu\right),$$

each of which obviously belongs to the ideal $\mathfrak{N}$ whenever f does, and since $\mathfrak{N}$ is closed, it is evident that $\mathfrak{N}$ has the desired property. Thus $\mathfrak{F}=0$ implies $\mathfrak{N}=\mathfrak{M}$. In the general case, we apply the results of Theorems 81 and 84 by considering the construction and properties of the space $\mathfrak{R}_{\mathfrak{N}}$. The sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ can be characterized as follows: if $\mathfrak{r}$ is in $\mathfrak{G}$, then $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ consists of the point $\mathfrak{r}$ alone; and, if $\mathfrak{r}$ is in $\mathfrak{F}$, then $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})=\mathfrak{F}$. First, if $\mathfrak{r}\epsilon\mathfrak{G}$, there exists a function f in $\mathfrak{N}$ with the property $f(\mathfrak{r})\neq 0$; and, remembering that f must vanish in $\mathfrak{F}$ and be constant in $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, we conclude that $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ is disjoint from $\mathfrak{F}$. Further-

more if $\mathfrak{s}$ is a point of $\mathfrak{S}$ distinct from $\mathfrak{r}$, there exists a function g in $\mathfrak{M}$ with the properties $g(\mathfrak{r})=1$, $g(\mathfrak{s})=0$. The function $h=fg$ belongs to the ideal $\mathfrak{N}$ together with f and has the properties $h(\mathfrak{r})\neq 0$, $h(\mathfrak{s})=0$. Remembering again that h must be constant in $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$, we conclude that $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ does not contain $\mathfrak{s}$. We see therefore that $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})=\{\mathfrak{r}\}$. Since the sets $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ are disjoint, it follows that $\mathfrak{r}\epsilon\mathfrak{F}$ implies $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})\subset\mathfrak{F}$. On the other hand, the fact that every function in $\mathfrak{N}$ vanishes in $\mathfrak{F}$ implies that $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})\supset\mathfrak{F}$ when $\mathfrak{r}\epsilon\mathfrak{F}$. Consequently, we have $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})=\mathfrak{F}$ when $\mathfrak{r}\epsilon\mathfrak{F}$, as we wished to prove. Theorem 84 now shows that the analytical subring $\mathfrak{N}^*$ generated by $\mathfrak{N}$ consists of those functions in $\mathfrak{M}$ which are constant on $\mathfrak{F}$: for such functions are precisely the ones which are constant on each set $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r})$ and hence remain single-valued on passing from f to $f_{\mathfrak{N}}$ in the manner described in that theorem. On the other hand, we can construct $\mathfrak{N}^*$ directly. Let $\mathfrak{P}$ be the class of all functions of the form $f+\alpha$ where $f\epsilon\mathfrak{N}$ and $\alpha\epsilon R$. It is evident that $\mathfrak{N}\subset\mathfrak{P}\subset\mathfrak{N}^*$; and also, by virtue of the fact that $0\epsilon\mathfrak{N}$, that $R\subset\mathfrak{P}$. If we can now show that $\mathfrak{P}$ is a closed subring of $\mathfrak{M}$, these relations enable us to conclude that it coincides with $\mathfrak{N}^*$: for the second identifies $\mathfrak{P}$ as an analytical subring; and the first then shows that $\mathfrak{P}=\mathfrak{N}^*$. To show that $\mathfrak{P}$ is closed, we recall that we are assuming $\mathfrak{F}$ to be non-void. If $\{f_n+\alpha_n\}$ is a convergent sequence in $\mathfrak{P}$, the fact that f_n vanishes in $\mathfrak{F}$ shows that the sequences $\{\alpha_n\}$ and $\{f_n\}$ converge separately. Since $\mathfrak{N}$ is a closed subset in the complete metric space $\mathfrak{M}$, $\{f_n\}$ has a limit in $\mathfrak{N}$. Thus $\mathfrak{P}$ is closed. Finally we show that $\mathfrak{P}$ is a subring: if $f+\alpha$ and $g+\beta$ are in $\mathfrak{P}$, then the difference $(f+\alpha)-(g+\beta)=(f-g)+(\alpha-\beta)$ and the product $(f+\alpha)(g+\beta)=(fg+f\beta+g\alpha)+(\alpha\beta)$ are in $\mathfrak{P}$ because the relations $f\epsilon\mathfrak{N}$, $g\epsilon\mathfrak{N}$, $\alpha\epsilon R$, $\beta\epsilon R$, and $R\subset\mathfrak{M}$ imply $f-g\epsilon\mathfrak{N}$, $fg+f\beta+g\alpha\epsilon\mathfrak{N}$, $\alpha-\beta\epsilon R$ and $\alpha\beta\epsilon R$. Since $\mathfrak{P}=\mathfrak{N}^*$, we now infer that every function in $\mathfrak{M}$ which is constant in $\mathfrak{F}$ has the form $f+\alpha$ where f is in $\mathfrak{N}$ and α is its value in $\mathfrak{F}$. In particular, we see that every function $\mathfrak{M}$ which vanishes in $\mathfrak{F}$ belongs to $\mathfrak{N}$. Thus the relation between the ideal $\mathfrak{N}$ and the corresponding closed set $\mathfrak{F}$ is that described in the theorem. The case where $\mathfrak{F}=0$, $\mathfrak{N}=\mathfrak{M}$ may obviously be included under this statement. On the other hand, we can easily verify that, if $\mathfrak{F}$ is an arbitrary closed set in $\mathfrak{R}$, the class of all functions in $\mathfrak{M}$ which vanish in $\mathfrak{F}$ is a closed ideal $\mathfrak{N}$. Thus the correspondence between closed ideals $\mathfrak{N}$ and closed sets $\mathfrak{F}$ is biunivocal.

In order to determine the nature of the quotient-ring $\mathfrak{M}/\mathfrak{N}$, where $\mathfrak{N}$ is a closed ideal, we consider the associated closed set $\mathfrak{F}$. We may discard the trivial case where $\mathfrak{F}=0$, $\mathfrak{N}=\mathfrak{M}$. When $\mathfrak{F}\neq 0$, we define the correspondence $f\rightarrow f_{\mathfrak{F}}$, where $f_{\mathfrak{F}}$ is obtained from a given function f in $\mathfrak{M}$ by restricting it to the closed set $\mathfrak{F}$. It is evident that this correspondence determines a homo-

morphism between $\mathfrak{M}$ and a subring of the function-ring for $\mathfrak{F}$. Since $f_{\mathfrak{F}}$ vanishes in $\mathfrak{F}$ if and only if f is in $\mathfrak{N}$, we see that this subring is an isomorph of $\mathfrak{M}/\mathfrak{N}$. Now if g is any continuous function in $\mathfrak{F}$, a fundamental theorem concerning normal spaces in general and bicompact H-spaces in particular shows that there exists a function f in $\mathfrak{M}$ such that $f_{\mathfrak{F}}=g$ in $\mathfrak{F}$.* Hence the indicated subring is identical with the function-ring for $\mathfrak{F}$; and this function-ring is an isomorph of $\mathfrak{M}/\mathfrak{N}$. We may remark that $\mathfrak{F}$, being closed in $\mathfrak{R}$, is a bicompact H-space.

The intersection or product of divisorless ideals is certainly an ideal; but, since the divisorless ideals in $\mathfrak{M}$ are closed in accordance with Theorem 75, it is also closed. On the other hand, if $\mathfrak{N}$ is a closed ideal distinct from $\mathfrak{M}$ and $\mathfrak{F}$ is the associated non-void closed set, we see that the divisorless ideals $\mathfrak{A}$ containing $\mathfrak{N}$ are precisely those determined by the points of $\mathfrak{F}$ in accordance with Theorem 80. Obviously, f belongs to the product of such divisorless ideals if and only if it vanishes in $\mathfrak{F}$; or, in other words, if and only if it belongs to $\mathfrak{N}$. Hence $\mathfrak{N}$ is the product of the divisorless ideals which contain it.

We pass now to the consideration of isomorphism and subrings. Here we merely summarize the results of Theorems 80–84 in somewhat different language.

THEOREM 86. *Two bicompact H-spaces are topologically equivalent if and only if their function-rings are analytically isomorphic. One bicompact H-space is a continuous image of another if and only if its function-ring is analytically isomorphic to an analytical subring of the function-ring of the other.*

Finally we shall state without proof an equivalent of Tychonoff's imbedding theorem for bicompact H-spaces.†

THEOREM 87. *Let* $\mathfrak{c}$ *be an arbitrary infinite cardinal number; let* A *be an arbitrary class of cardinal number* $\mathfrak{c}$, *for example, the class of all ordinal numbers preceding some suitable (even the first suitable) ordinal number* ω; *let* $\mathfrak{R}_{\mathfrak{c}}$ *be the class of all real functions* $\mathfrak{r}=\mathfrak{r}(\alpha)$ *defined over* A, *where* $0\leqq\mathfrak{r}(\alpha)\leqq 1$; *and let* $B_{\mathfrak{c}}$ *be the class of all sets in* $\mathfrak{R}_{\mathfrak{c}}$ *generated from the special sets* $\mathfrak{W}_{\alpha}$, *specified by the inequalities* $\rho<\mathfrak{r}(\alpha)<\sigma$ *where* ρ *and* σ *are rational numbers, by the formation of finite intersections. By the assignment of each non-void set belonging to* $B_{\mathfrak{c}}$ *as a neighborhood of every one of its points,* $\mathfrak{R}_{\mathfrak{c}}$ *becomes a bicompact H-space of character* $\mathfrak{c}$. *Every bicompact H-space of character not exceeding* $\mathfrak{c}$ *is topologically equivalent to a closed subset of* $\mathfrak{R}_{\mathfrak{c}}$; *and its function-ring is a homomorph of the function-ring of* $\mathfrak{R}_{\mathfrak{c}}$, *in the sense indicated in Theorem* 85.

* AH, pp. 73–76.

† Tychonoff, Mathematische Annalen, vol. 102 (1930), pp. 544–561.

We point out the rather striking parallel between the results obtained in Theorems 85–87 for bicompact H-spaces and those obtained in Theorems 4, 7, 9, 10 for Boolean spaces. From this observation, we may surmise that both groups of theorems have the same, essentially algebraic, origin. To discover this common origin, if there be any, it would apparently be necessary to give an abstract characterization of function-rings.

We turn now to the application of the theory of function-rings to problems in the theory of extensions. Some of our results bring out the quite remarkable properties of the bicompact H-extension of a CR-space which we have already constructed in Theorems 78 and 79: for we can generalize the latter theorem in a quite complete way. We shall now state this generalization.

THEOREM 88. *Let $\mathfrak{R}$ be a CR-space; let $\mathfrak{Q}$ be the bicompact strict H-extension of $\mathfrak{R}$ constructed in Theorems* 78 *and* 79; *let $\mathfrak{T}$ be a CR-space which is a continuous image of $\mathfrak{R}$ by virtue of a correspondence* $\mathfrak{t}=\tau(\mathfrak{r})$; *and let $\mathfrak{S}$ be any bicompact immediate or strict H-extension of $\mathfrak{T}$. Then there exists a continuous univocal correspondence $\mathfrak{s}=\sigma(\mathfrak{q})$ from $\mathfrak{Q}$ to $\mathfrak{S}$ which coincides in $\mathfrak{R}$ with $\tau(\mathfrak{r})$. In particular, every bicompact immediate or strict H-extension of $\mathfrak{R}$ is a continuous image of $\mathfrak{Q}$.*

If $f_{\mathfrak{S}}$ is any function in the function-ring for $\mathfrak{S}$, the replacement of its argument by $\tau(\mathfrak{r})$ yields a function $f_{\mathfrak{S}}(\tau(\mathfrak{r}))$ in the function-ring for $\mathfrak{R}$. By Theorem 79, the latter function can be extended in a unique way over the space $\mathfrak{Q}$ so as to yield a function $f_{\mathfrak{Q}}$ in the function-ring for $\mathfrak{Q}$. By virtue of the fact that $\mathfrak{R}$ and $\mathfrak{T}$ are everywhere dense in $\mathfrak{Q}$ and in $\mathfrak{S}$ respectively, the correspondence $f_{\mathfrak{S}}\rightarrow f_{\mathfrak{Q}}$ is seen to be an analytical isomorphism between the function-ring for $\mathfrak{S}$ and a certain analytical subring of the function-ring for $\mathfrak{Q}$. By Theorem 86 the space $\mathfrak{S}$ is a continuous image of $\mathfrak{Q}$. We examine the relation between $\mathfrak{S}$ and $\mathfrak{Q}$ in greater detail. If $\mathfrak{N}$ is the analytical subring consisting of the functions $f_{\mathfrak{Q}}$, we construct the space $\mathfrak{Q}_{\mathfrak{N}}$ described in Theorem 81. The correspondence $f_{\mathfrak{Q}}\rightarrow f_{\mathfrak{N}}$ defined by $f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{q}))=f_{\mathfrak{Q}}(\mathfrak{q})$, $\mathfrak{q}\epsilon\mathfrak{Q}$, is an analytical isomorphism between $\mathfrak{N}$ and the function-ring for $\mathfrak{Q}_{\mathfrak{N}}$ in accordance with Theorems 81 and 84. The function-rings for $\mathfrak{S}$ and $\mathfrak{Q}_{\mathfrak{N}}$, being analytically isomorphic to $\mathfrak{N}$, are analytically isomorphic to each other. The correspondence $f_{\mathfrak{N}}\rightarrow f_{\mathfrak{S}}$ therefore defines a topological equivalence $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{q})\longleftrightarrow\mathfrak{s}$ between $\mathfrak{Q}_{\mathfrak{N}}$ and $\mathfrak{S}$ as described in detail in Theorem 83. Thus the continuous correspondence from $\mathfrak{Q}$ to $\mathfrak{S}$ is obtained by eliminating $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{q})$ from the correspondences $\mathfrak{q}\longleftrightarrow\mathfrak{X}_{\mathfrak{N}}(\mathfrak{q})$, $\mathfrak{X}_{\mathfrak{N}}(\mathfrak{q})\longleftrightarrow\mathfrak{s}$. For a point $\mathfrak{r}$ in $\mathfrak{R}$, we have $f_{\mathfrak{N}}(\mathfrak{X}_{\mathfrak{N}}(\mathfrak{r}))=f_{\mathfrak{Q}}(\mathfrak{r})=f_{\mathfrak{S}}(\tau(\mathfrak{r}))$. Hence we see that the correspondence from $\mathfrak{Q}$ to $\mathfrak{S}$ carries $\mathfrak{r}$ into $\tau(\mathfrak{r})$, as we wished to prove.

If we take $\mathfrak{T}=\mathfrak{R}$, we see that $\mathfrak{S}$ is a continuous image of $\mathfrak{Q}$, as before.

Hence every bicompact immediate or strict H-extension $\mathfrak{S}$ of $\mathfrak{R}$ is a continuous image of $\mathfrak{Q}$.

By a quite similar argument we can now complete the information obtained in Theorem 77.

THEOREM 89. *If a CR-space $\mathfrak{T}$ is a continuous image of a T_0-space $\mathfrak{R}$, then $\mathfrak{T}$ is also a continuous image of the associated CR-space $\mathfrak{R}^*$ constructed in Theorem 77.*

If f is any function in the function-ring for $\mathfrak{T}$ and if τ is the continuous correspondence from $\mathfrak{R}$ to $\mathfrak{T}$, then $f(\tau(\mathfrak{r}))$ is in the function-ring for $\mathfrak{R}$; and the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$ from $\mathfrak{R}$ to $\mathfrak{R}^*$ carries it into a function f^* in the function-ring for $\mathfrak{R}^*$. If we place $\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{t}$ in correspondence whenever $\mathfrak{r}\rightarrow\mathfrak{X}(\mathfrak{r})$ and $\mathfrak{r}\rightarrow\mathfrak{t}=\tau(\mathfrak{r})$, we see therefore that the antecedent of the set $\mathfrak{T}(\alpha<f<\beta)$ is the set $\mathfrak{R}^*(\alpha<f^*<\beta)$, in accordance with the relation $f^*(\mathfrak{X}(\mathfrak{r}))=f(\tau(\mathfrak{r}))$. If $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are distinct points in $\mathfrak{T}$, there exists a function f which belongs to the function-ring for $\mathfrak{T}$ and has the properties $f(\mathfrak{t}_1)=0$, $f(\mathfrak{t}_2)=1$. Since the sets $\mathfrak{T}(-1/2<f<1/2)$, $\mathfrak{T}(1/2<f<3/2)$ are disjoint, their antecedents $\mathfrak{R}^*(-1/2<f^*<1/2)$, $\mathfrak{R}^*(1/2<f^*<3/2)$ are likewise disjoint. It follows that the correspondence $\mathfrak{X}(\mathfrak{r})\rightarrow\mathfrak{t}$ defined above is univocal. Since the sets $\mathfrak{T}(\alpha<f<\beta)$ constitute a basis for $\mathfrak{T}$, and since their antecedents $\mathfrak{R}^*(\alpha<f^*<\beta)$ are open, the correspondence is also continuous. Hence this correspondence represents $\mathfrak{T}$ as a continuous image of $\mathfrak{R}^*$ as well as of $\mathfrak{R}$.

We shall now consider in some detail the nature of the bicompact immediate H-extensions of a CR-space and their connections with the mapping theory. First, let us consider a modification of Theorems 78 and 79.

THEOREM 90. *If $\mathfrak{R}$ is a CR-space of infinite character $\mathfrak{c}$, then $\mathfrak{R}$ has a bicompact immediate, and hence strict, H-extension $\mathfrak{Q}$ of the same character $\mathfrak{c}$.*

In the given space $\mathfrak{R}$, there exists a basis of cardinal number $\mathfrak{c}$ consisting of sets $\mathfrak{R}(\alpha<f<\beta)$ where f belongs to the function-ring $\mathfrak{M}$ for $\mathfrak{R}$. The associated functions f constitute a subclass $\mathfrak{N}$ of $\mathfrak{M}$ with cardinal number not exceeding $\mathfrak{c}$. If $\mathfrak{Q}$ is the bicompact immediate H-extension of $\mathfrak{R}$ constructed in Theorems 78 and 79, every function in $\mathfrak{M}$ can be extended from $\mathfrak{R}$ to $\mathfrak{Q}$ in accordance with Theorem 79. By this extension, $\mathfrak{M}$ and $\mathfrak{N}$ are replaced by $\mathfrak{M}_{\mathfrak{Q}}$, the function-ring for $\mathfrak{Q}$, and a subclass $\mathfrak{P}=\mathfrak{N}_{\mathfrak{Q}}$ respectively. The construction of Theorem 81 yields a bicompact H-space $\mathfrak{Q}_{\mathfrak{P}}$ which is a continuous image of $\mathfrak{Q}$ by virtue of a correspondence $\mathfrak{q}\rightarrow\mathfrak{X}_{\mathfrak{P}}(\mathfrak{q})$. This correspondence carries $\mathfrak{P}$ into a subclass $\mathfrak{N}_{\mathfrak{P}}$ of the function-ring $\mathfrak{M}_{\mathfrak{P}}$ for $\mathfrak{Q}_{\mathfrak{P}}$. The analytical subrings generated by $\mathfrak{P}$ and $\mathfrak{N}_{\mathfrak{P}}$ respectively are analytically isomorphic under this correspondence; and Theorem 84 shows that the analytical subring

generated by $\mathfrak{N}_{\mathfrak{P}}$ coincides with $\mathfrak{M}_{\mathfrak{P}}$. If we refer to Theorem 82, we see that the sets $\mathfrak{Q}_{\mathfrak{P}}(\alpha<f_{\mathfrak{P}}<\beta)$, $f_{\mathfrak{P}}\epsilon\mathfrak{N}_{\mathfrak{P}}$, where α and β are unrestricted real numbers, constitute a basis for $\mathfrak{Q}_{\mathfrak{P}}$. Obviously, we may restrict α and β to be rational without disturbing the stated property. Since the cardinal number of $\mathfrak{N}_{\mathfrak{P}}$ does not exceed $\mathfrak{c}$, we infer that the character of $\mathfrak{Q}_{\mathfrak{P}}$ does not exceed $\mathfrak{c}$. It is evident that the correspondence from $\mathfrak{Q}$ to $\mathfrak{Q}_{\mathfrak{P}}$ carries $\mathfrak{R}$, which is an everywhere dense subset of $\mathfrak{Q}$, into a set $\mathfrak{R}_{\mathfrak{P}}$ everywhere dense in $\mathfrak{Q}_{\mathfrak{P}}$. Since the correspondence $\mathfrak{q}\rightarrow\mathfrak{X}_{\mathfrak{P}}(\mathfrak{q})$ carries $f(\mathfrak{q})$ into $f_{\mathfrak{P}}(\mathfrak{X}_{\mathfrak{P}}(\mathfrak{q}))$, we see that $\mathfrak{q}=\mathfrak{r}\epsilon\mathfrak{R}$ implies $f(\mathfrak{r})=f_{\mathfrak{P}}(\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r}))$, $\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r})\epsilon\mathfrak{R}_{\mathfrak{P}}$. Hence the sets $\mathfrak{R}(\alpha<f<\beta)$ are carried into the sets $\mathfrak{R}_{\mathfrak{P}}(\alpha<f_{\mathfrak{P}}<\beta)$ whenever $f\epsilon\mathfrak{N}$ and $f_{\mathfrak{P}}\epsilon\mathfrak{N}_{\mathfrak{P}}$. We infer that the correspondence $\mathfrak{r}\rightarrow\mathfrak{X}_{\mathfrak{P}}(\mathfrak{r})$ carries $\mathfrak{R}$ into $\mathfrak{R}_{\mathfrak{P}}$ biunivocally and bicontinuously. Hence $\mathfrak{R}$ and $\mathfrak{R}_{\mathfrak{P}}$ are topologically equivalent. Since the character of $\mathfrak{R}_{\mathfrak{P}}$ is $\mathfrak{c}$, the character of $\mathfrak{Q}_{\mathfrak{P}}$ is not less than $\mathfrak{c}$. Thus $\mathfrak{Q}_{\mathfrak{P}}$ is the desired extension.

We now return to the direct study of Boolean maps with particular reference to CR-spaces.

THEOREM 91. *If $\mathfrak{R}$ is a T_0-space with a Boolean map $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ in which $\mathfrak{X}$ is a subfamily of a continuous covering family $\mathfrak{Z}$, then it is a CR-space.*

By Theorem 22, the space $\mathfrak{Q}$ defined by topologizing the family $\mathfrak{Z}$ is a continuous image of $\mathfrak{B}$; and $\mathfrak{Q}$ is therefore a bicompact H-space in accordance with Theorem 72. Since $\mathfrak{Q}$ contains $\mathfrak{R}$ as a subspace, it follows that $\mathfrak{R}$ is a CR-space.

THEOREM 92. *In a CR-space $\mathfrak{R}$, let $\mathcal{G}$ be a basis consisting of sets $\mathfrak{R}(\alpha<f<\beta)$ where, for each f, the corresponding real numbers α and β are allowed to range over sets everywhere dense in R; and let A be the basic ring for $\mathfrak{R}$ generated by $\mathcal{G}$. If $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ is the algebraic map defined by A, then there exists a continuous covering family $\mathfrak{Z}$ in $\mathfrak{B}$ which contains $\mathfrak{X}$ as a subfamily. The topological space $\mathfrak{Q}$ defined by $\mathfrak{Z}$ is a bicompact immediate H-extension of $\mathfrak{R}$. Similarly, if $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ is the irredundant (algebraic) map generated by A through the processes described in Theorems 61–63, the family $\mathfrak{X}$ can again be imbedded as a subfamily in a continuous covering family $\mathfrak{Z}$.*

We start the proof with a construction like that employed in Theorem 90, using the class $\mathfrak{N}$ of all functions f occurring in the given basis-sets $\mathfrak{R}(\alpha<f<\beta)$. Thus we imbed $\mathfrak{R}$ as an everywhere dense set in a bicompact H-space $\mathfrak{Q}_{\mathfrak{P}}$, in such a way that every function f in $\mathfrak{N}$ can be extended continuously from $\mathfrak{R}$ to $\mathfrak{Q}_{\mathfrak{P}}$ as a function $f_{\mathfrak{P}}$. The sets $\mathfrak{Q}_{\mathfrak{P}}(\alpha<f_{\mathfrak{P}}<\beta)$ where $f\epsilon\mathfrak{N}$ then constitute a basis for $\mathfrak{Q}_{\mathfrak{P}}$; and this statement remains true even if the numbers α and β corresponding to each function $f_{\mathfrak{P}}$ are restricted to everywhere dense sets in R. Thus each set $\mathfrak{R}(\alpha<f<\beta)$ determines a corresponding

set $\mathfrak{Q}_{\mathfrak{P}}(\alpha<f_{\mathfrak{P}}<\beta)$, the latter sets constituting a basis $\mathcal{G}_{\mathfrak{P}}$ in $\mathfrak{Q}_{\mathfrak{P}}$. Since $\mathfrak{R}\mathfrak{Q}_{\mathfrak{P}}(\alpha<f_{\mathfrak{P}}<\beta)=\mathfrak{R}(\alpha<f<\beta)$, we see that the basic ring $A_{\mathfrak{P}}$ for $\mathfrak{Q}_{\mathfrak{P}}$ generated by the basis $\mathcal{G}_{\mathfrak{P}}$ is related to A in the manner described in Theorem 39 for the rings A and B respectively. Hence we can apply Theorem 40 to the study of the maps generated by $A_{\mathfrak{P}}$ and A. Since $\mathfrak{R}$ is a CR- and hence an SR-space, $\mathfrak{Q}_{\mathfrak{P}}$ is a strict H-extension of $\mathfrak{R}$ in accordance with Theorem 65. Thus we see that the map generated by $A_{\mathfrak{P}}$ can be simplified by the removal of the set $\mathfrak{E}(\mathfrak{b})$ corresponding to the ideal $\mathfrak{b}$ of those members of $A_{\mathfrak{P}}$ which are nowhere dense subsets of $\mathfrak{R}'$. The resulting map is equivalent to one obtained from $m(\mathfrak{R}, \mathfrak{B}, \mathfrak{X})$ by augmenting the family $\mathfrak{X}$. We thus have a map $m(\mathfrak{Q}_{\mathfrak{P}}, \mathfrak{B}, \mathfrak{Z})$. In order to prove the present theorem, it is thus sufficient to show that $\mathfrak{Z}$ is a continuous covering family. We know that the sets in the family $\mathfrak{Z}$ are disjoint by virtue of the fact that $\mathfrak{Q}_{\mathfrak{P}}$ is an H-space. Using the bicompactness of $\mathfrak{Q}_{\mathfrak{P}}$, we can prove that $\mathfrak{Z}$ covers $\mathfrak{B}$. Evidently, we can return to the original map $m(\mathfrak{Q}_{\mathfrak{P}}, \mathfrak{E}(A_{\mathfrak{P}}), \mathfrak{Y})$ and prove instead that $\mathfrak{Y}$ covers $\mathfrak{E}(A_{\mathfrak{P}})$. If there exists a point $\mathfrak{p}$ in $\mathfrak{E}(A_{\mathfrak{P}})$ which belongs to no set in $\mathfrak{Y}$, we can construct for each set $\mathfrak{Y}$ in $\mathfrak{Y}$ a set $\mathfrak{E}(a)$, $a\epsilon A_{\mathfrak{P}}$, such that $\mathfrak{Y}\subset\mathfrak{E}(a)$, $\mathfrak{p}\epsilon\mathfrak{E}'(a)$. Since the elements a thus obtained constitute a family of subsets of $\mathfrak{Q}_{\mathfrak{P}}$ with the property that their interiors cover $\mathfrak{Q}_{\mathfrak{P}}$, we see that it is possible to select among them certain ones $a_1, \cdots, a_n$, such that $a_1\vee\cdots\vee a_n>a_1'^{-\prime}\vee\cdots\vee a_n'^{-\prime}>\mathfrak{Q}_{\mathfrak{P}}$. Since the relation $a_1\vee\cdots\vee a_n=\mathfrak{Q}_{\mathfrak{P}}$ implies $\mathfrak{p}\epsilon\mathfrak{E}(A_{\mathfrak{P}})=\mathfrak{E}(a_1\vee\cdots\vee a_n)=\mathfrak{E}(a_1)\cup\cdots\cup\mathfrak{E}(a_n)$, we reach a contradiction. Thus $\mathfrak{Y}$ covers $\mathfrak{E}(A_{\mathfrak{P}})$, and $\mathfrak{Z}$ covers $\mathfrak{B}$. Using the fact that $\mathfrak{Q}_{\mathfrak{P}}$ is necessarily an R-space in accordance with Theorem 71, we can now prove that the family $\mathfrak{Z}$ is continuous. Since $A_{\mathfrak{P}}$ contains the basis $\mathcal{G}_{\mathfrak{P}}$ for the R-space $\mathfrak{Q}_{\mathfrak{P}}$, Theorem 69 shows that in the map $m(\mathfrak{Q}_{\mathfrak{P}}, \mathfrak{E}(A_{\mathfrak{P}}), \mathfrak{Y})$ the family $\mathfrak{Y}$ is continuous. Accordingly, $\mathfrak{Q}_{\mathfrak{P}}$ is a continuous image of the bicompact Boolean space $\mathfrak{E}(A_{\mathfrak{P}})$ as we see by reference to Theorem 22. It follows that the removal of the set $\mathfrak{E}(\mathfrak{b})$ from $\mathfrak{E}(A_{\mathfrak{P}})$ leaves $\mathfrak{Q}_{\mathfrak{P}}$ a continuous image of the set $\mathfrak{E}'(\mathfrak{b})$. Hence the map $m(\mathfrak{Q}_{\mathfrak{P}}, \mathfrak{B}, \mathfrak{Z})$ represents $\mathfrak{Q}_{\mathfrak{P}}$ as a continuous image of $\mathfrak{B}$. The family $\mathfrak{Z}$ is therefore continuous in accordance with Theorem 22.

Since $\mathfrak{Q}_{\mathfrak{P}}$ and $\mathfrak{R}$ are both CR-spaces, we can also remove from $\mathfrak{B}=\mathfrak{E}(A)$ the set $\mathfrak{E}(\mathfrak{a})$ corresponding to the ideal $\mathfrak{a}$ of all nowhere dense sets in A, as described in Theorem 65. We thus obtain maps $m(\mathfrak{Q}_{\mathfrak{P}},\mathfrak{E}'(\mathfrak{a}),\mathfrak{Z})$ and $m(\mathfrak{R}, \mathfrak{E}'(\mathfrak{a}), \mathfrak{X})$ where $\mathfrak{Z}$ is an extension of the family $\mathfrak{X}$. It is evident that $\mathfrak{Z}$ covers $\mathfrak{E}'(\mathfrak{a})$; and the argument used above can be applied again to show that $\mathfrak{Z}$ is continuous.

Theorems 91 and 92 show that the CR-spaces are precisely those which have certain Boolean maps with the property that the family $\mathfrak{X}$ can be ex-

tended to yield a continuous covering family $\mathcal{Z}$. In order to complete the theory, we consider the problem of constructing such an extension in an arbitrary Boolean map.

THEOREM 93. *If $m(\mathfrak{R}, \mathfrak{B}, \mathcal{X})$ is an arbitrary Boolean map of a T_0-space $\mathfrak{R}$ and if $\mathfrak{M}$ is the function-ring for $\mathfrak{B}$, then the class $\mathfrak{N}$ of all functions in $\mathfrak{M}$ which are constant in each set of the family $\mathcal{X}$ is an analytical subring of $\mathfrak{M}$. The space $\mathfrak{B}_{\mathfrak{N}}$ constructed by the processes described in Theorem* 81 *is a continuous image of $\mathfrak{B}$ defined by the Boolean map $m(\mathfrak{B}_{\mathfrak{N}}, \mathfrak{B}, \mathcal{X}_{\mathfrak{N}})$. Each set $\mathfrak{X}$ in $\mathcal{X}$ is contained in a unique set $\mathfrak{X}_{\mathfrak{N}}$ in $\mathcal{X}_{\mathfrak{N}}$; and the correspondences $\mathfrak{r} \rightarrow \mathfrak{X}(\mathfrak{r}) \rightarrow \mathfrak{X}_{\mathfrak{N}}$, where $\mathfrak{X}(\mathfrak{r}) \subset \mathfrak{X}_{\mathfrak{N}}$, determine a continuous image of $\mathfrak{R}$ in $\mathfrak{B}_{\mathfrak{N}}$. In order that there exist a continuous covering family $\mathcal{Z}$ in $\mathfrak{B}$ which contains the family $\mathcal{X}$ as a subfamily, it is necessary and sufficient that $\mathfrak{X} \subset \mathfrak{X}_{\mathfrak{N}}$ imply $\mathfrak{X} = \mathfrak{X}_{\mathfrak{N}}$; when this condition is satisfied, the family $\mathcal{Z}$ may be taken as $\mathcal{Z} = \mathcal{X}_{\mathfrak{N}}$.*

The class $\mathfrak{N}$ can be constructed in the following way. If $\mathfrak{X}$ is any set in $\mathcal{X}$, it is closed. As in the proof of Theorem 85, we see that the class $\mathfrak{N}(\mathfrak{X})$ of all functions in $\mathfrak{M}$ which are constant in $\mathfrak{X}$ is an analytical subring. It is obvious that $\mathfrak{N}$ is the intersection of all the analytical subrings $\mathfrak{N}(\mathfrak{X})$. Hence $\mathfrak{N}$ is also an analytical subring; that is, it is a closed subring which contains every constant function. We can now construct $\mathfrak{X}_{\mathfrak{N}} = \mathfrak{X}_{\mathfrak{N}}(\mathfrak{p})$, $\mathcal{X}_{\mathfrak{N}}$, and $\mathfrak{B}_{\mathfrak{N}}$ as described in Theorem 81. It is evident from the construction of $\mathfrak{N}$ that $\mathfrak{p} \in \mathfrak{X}$ implies $\mathfrak{X} \subset \mathfrak{X}_{\mathfrak{N}}(\mathfrak{p})$. In Theorem 81, we saw that the correspondence $\mathfrak{p} \rightarrow \mathfrak{X}_{\mathfrak{N}}(\mathfrak{p})$ from $\mathfrak{B}$ to $\mathfrak{B}_{\mathfrak{N}}$ is univocal and continuous; and Theorem 22 shows that this correspondence defines a Boolean map $m(\mathfrak{B}_{\mathfrak{N}}, \mathfrak{B}, \mathcal{X}_{\mathfrak{N}})$. If we place $\mathfrak{X}_{\mathfrak{N}}$ in correspondence with $\mathfrak{X}$ whenever $\mathfrak{X} \subset \mathfrak{X}_{\mathfrak{N}}$, we obtain a univocal correspondence carrying $\mathfrak{R}$ into a subspace of $\mathfrak{B}_{\mathfrak{N}}$; we identify $\mathfrak{R}$ with the family $\mathcal{X}$, topologized in the usual way, of course. This correspondence can also be obtained by restricting the continuous correspondence σ, from $\mathfrak{B}$ to $\mathfrak{B}_{\mathfrak{N}}$, to the union $\mathfrak{S}(\mathcal{X})$ of all the sets in $\mathcal{X}$. This description of the correspondence shows immediately that it is continuous.

From the results already obtained, it is clear that the family $\mathcal{X}_{\mathfrak{N}}$ is a continuous covering family. Hence, if $\mathfrak{X} \subset \mathfrak{X}_{\mathfrak{N}}$ implies $\mathfrak{X} = \mathfrak{X}_{\mathfrak{N}}$, we can obtain a continuous covering family $\mathcal{Z}$ which contains $\mathcal{X}$ merely by taking $\mathcal{Z} = \mathfrak{X}_{\mathfrak{N}}$. On the other hand, if $\mathcal{X}$ has such an extension $\mathcal{Z}$, we impose upon $\mathcal{Z}$ the usual topology, obtaining a bicompact H-space $\mathfrak{Q}$: for Theorem 22 shows that $\mathfrak{Q}$ is a continuous image of $\mathfrak{B}$; and Theorem 72 then shows that $\mathfrak{Q}$ is a bicompact H-space. If ρ is the correspondence from $\mathfrak{B}$ to $\mathfrak{Q}$ and if f is any function in the function-ring for $\mathfrak{Q}$, then $f(\rho(\mathfrak{p}))$ is a bounded continuous function in $\mathfrak{B}$ which is constant in each set $\mathfrak{Z}$. Since, in particular, $f(\rho(\mathfrak{p}))$ is constant in each set $\mathfrak{X}$, it belongs to the class $\mathfrak{N}$ discussed above. Since $\mathfrak{Q}$ is a bicompact

H-space, it is a CR-space. Hence f can be chosen so as to assume distinct values in any two distinct points which we may prescribe. Consequently, the corresponding function $f(\rho(\mathfrak{p}))$ may be chosen so as to assume distinct values in any two distinct sets in $\mathfrak{Z}$ which we may prescribe. It follows that each set $\mathfrak{Z}$ contains a set $\mathfrak{X}_{\mathfrak{R}}$. In particular, we infer that $\mathfrak{X} \subset \mathfrak{X}_{\mathfrak{R}}$ implies $\mathfrak{X} = \mathfrak{X}_{\mathfrak{R}}$, as we wished to show.

Harvard University,
Cambridge, Mass.

ERRATA, VOLUME 41

A. P. Morse, *Convergence in variation and related topics*, pp. 48–83.

Page 55, line 4. Instead of $f'(t)=\mu$ read $|f'(t)|=\mu$.

Page 55. In the fourth line of the four line display replace $-$ by $+$.

Page 62, Lemma 4.4. Property (iii) should read, *if ϕ is any function in CC which is applicable to f, then ϕ is applicable to g and* $\cdots$.

Also in Lemma 4.4, the last member of the first line which defines $\Lambda(t)$ should read $|\phi:f(t)-\phi:f(t-)|$.

Page 64. In the statement of Lemma 4.5 and Corollary 4.1, replace g by f.

Page 74. In the second line of the statement of Theorem 5.8, replace $[0, 1]$ by $[0, t]$.

Page 80. The displayed line in the statement of Theorem 8.2 should be the same as the corresponding line in Theorem 8.3.

Oystein Ore, *On the theorem of Jordan-Hölder*, pp. 266–275.

In line 1, p. 274, there is a disturbing misprint. Theorem 7 is a consequence of Theorem 1, not of Theorem 5 as stated. It is essential that the conditions for Theorem 5 are *not* satisfied in this case. A simple calculation shows however that the conditions for Theorem 1 are fulfilled.